REPORT RESUMES

ED 003 178 56
COMPARATIVE RESEARCH ON METHODS AND MEDIA FOR PRESENTING PROGRAMED COURSES IN MATHEMATICS AND ENGLISH.
BY- CARPENTER, C.R. GREENHILL, L.P.
PENNSYLVANIA STATE UNIV., UNIVERSITY PARK
REPORT NUMBER NDEA-VIIA-567 PUB DATE MAR 63
GRANT OEG-736116
EDRS PRICE MF-$0.18 HC-$3.36 84P.

DESCRIPTORS- *PROGRAMED INSTRUCTION, *MATHEMATICS INSTRUCTION, *ENGLISH INSTRUCTION, COMPARATIVE ANALYSIS, *METHODS RESEARCH, *MEDIA RESEARCH, CLOSED CIRCUIT TELEVISION, PROGRAMED MATERIALS, PACING, TIME FACTORS (LEARNING), GROUP INSTRUCTION, UNIVERSITY PARK, PENNSYLVANIA

THE GENERAL PURPOSES WERE (1) TO INVESTIGATE THE POSSIBILITIES OF PROGRAMING ENTIRE COURSES IN MATHEMATICS AND ENGLISH GRAMMAR AND (2) TO COMPARE DIFFERENT METHODS AND MEDIA, WITH SPECIAL EMPHASIS ON CLOSED-CIRCUIT TELEVISION, FOR PRESENTING PROGRAMED LEARNING MATERIALS. TO ACCOMPLISH THE OBJECTIVES, IT WAS NECESSARY TO ANALYZE THE PROBLEMS OF PACING AND INDIVIDUAL AND GROUP CONDITIONS FOR THE LEARNING OF PROGRAMED COURSES. THIS RESEARCH QUESTIONED THAT THE BEST METHODS OF USING PROGRAMED MATERIALS ARE SELF-PACING AND INDIVIDUALIZED STUDY. IT WAS SUGGESTED THAT A LEARNER'S OWN SPEED MAY NOT BE THE OPTIMUM RATE. THE RESEARCH SHOWED SOME OF THE POSSIBILITIES OF USING COMBINATIONS OF NEW MEDIA WITH PACED PROGRAMS FOR INSTRUCTING LARGE NUMBERS OF STUDENTS SIMULTANEOUSLY IN TIME-LIMITED PERIODS OF INSTRUCTION. (LP)

ED003178

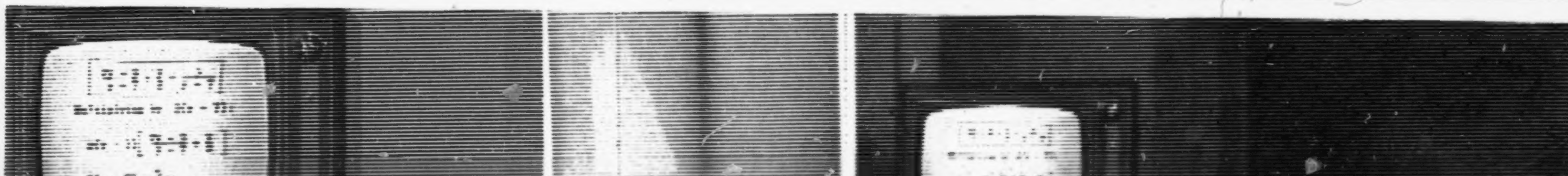
567

567

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

COMPARATIVE RESEARCH ON METHODS AND MEDIA FOR PRESENTING PROGRAMED COURSES IN MATHEMATICS AND ENGLISH

The Pennsylvania State University · University Park, Pennsylvania



COMPARATIVE RESEARCH ON METHODS AND MEDIA FOR PRESENTING PROGRAMED COURSES IN MATHEMATICS AND ENGLISH

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

a Cooperative Project by

The University Division of Instructional Services

and

The Departments of Mathematics and English

at The Pennsylvania State University

under the direction of

C. R. Carpenter and L. P. Greenhill

The research reported herein was supported by Grant No. 736116 from the United States Office of Education, Department of Health, Education, and Welfare, under Title VII of the National Defense Education Act 1958: "Research to Discover Methods of Applying and Testing Teaching Machine Principles of Programing Basic College Courses in Mathematics and English Composition for Presentation over Closed-Circuit Television."

The Pennsylvania State University
University Park, Pennsylvania
March 1963

CONTENTS

ACKNOWLEDGMENTS vi
SUMMARY vii

PART I

THE "NEW" MEDIA AND PROGRAMED LEARNING

Introduction 1
General Assumptions on Programing Procedures 1
Implications of the General Assumptions for Functions of Teaching Machines 4
Neglect of "New" Media for Programed Learning 4
Functional Requirements of Devices for Presentation of Programed Learning 5
Characteristics of "New" Media Relative to Programed Learning 6
Adapting Television and Programed Courses 7
Research Strategy 8
Summary 9
References 9

PART II

RESEARCH ON METHODS AND MEDIA FOR PRESENTING PROGRAMED MATHEMATICS

The Development of the Mathematics Program 11
Experiment I: Comparison of External-pacing and Self-pacing of Programed Instruction in Mathematics Using Different Methods of Presentation 13
  Purpose 13
  Statement of the Problem 13
  Method 14
  Results 20
  Discussion of Results 24
  Conclusions 25

Experiment II: An Experimental Comparison of Differential Rates of Pacing Programed Mathematics . . . 25
Purpose . . . 25
Statement of the Problem . . . 25
Method . . . 25
Results . . . 26
Discussion of Results . . . 28
Conclusions . . . 28
Experiment III: Adaptation and Evaluation of a Programed Mathematics Course for Televised Presentation . . . 29
Purpose . . . 29
Statement of the Problem . . . 29
Adaptation of the Programed Course for Television Presentation . . . 29
Effectiveness of Televised Presentations . . . 35
Method . . . 35
Results . . . 36
Discussion of Results and Conclusions . . . 38
Experiment IV: Paired vs. Individual Study of Programed Instruction in Contemporary Algebra . . . 39
Background and Purpose . . . 39
Method . . . 40
Results . . . 42
Discussion of Results and Conclusions . . . 45

PART III

RESEARCH ON METHODS AND MEDIA FOR PRESENTING PROGRAMED ENGLISH GRAMMAR

Development of the English Program . . . 47
Adaptation for Televised Presentation . . . 50
Method of Televising Program . . . 53
The Tests . . . 55
The Pilot Project: Fall, 1961 . . . 57
Experiment V: Comparisons of Televised with Teaching Machine and Televised with Instructor Presentations of English Grammar . . . 57
Objectives . . . 57

Method 57
Differences between the English and the Mathematics Experiments 59
Results 61
Discussion of Results and Conclusions 62
Experiment VI: Effects of Personality-Pairing on the Performance of Students in a Programed Course in English Grammar 63
Purpose 63
Method 64
Results 66
Discussion of Results and Conclusions 70

PART IV

OBSERVATIONS, JUDGMENTS, AND SUGGESTIONS FOR FURTHER RESEARCH

Observations and Judgments 73
Suggested Research and Development Work Based on these Studies 74

ACKNOWLEDGMENTS

The research reported herein is the result of a cooperative effort among many individuals representing the University Division of Instructional Services (formerly the Division of Academic Research and Services) and the Departments of Mathematics and English at The Pennsylvania State University.

At appropriate points throughout the report the contributions of the individuals participating in specific studies have been indicated. However, special mention should be made of the major contributors to these studies:

(1) For the Mathematics Department—James B. Bartoo, Department Head; Frank Kocher, Albert R. Erskine, and Hilda F. Duncan, Instructors in Mathematics; Michael Varano and Ivan Doszpoly, Graduate Assistants

(2) For the English Department—Henry W. Sams, Department Head; John S. Bowman, Professor of English Composition; Agnes R. McElwee, Professor of English Composition; and James P. Stewart, Associate Professor of English Composition

(3) For the University Division of Instructional Services—John Quackenbush, Project Coordinator; Ralph Heimer, Mathematics Program Writer; Ann Funk, English Program Writer; Richard E. Spencer, Data Analyst; Anthony Pagano and Walter Dick, Graduate Assistants; Luther F. Kepler, Jr., James Vitale, Ronald Slawson and the Instructional Television Staff; Joseph Servello, Graphics Artist; Edmond L. Seguin, Frank Simon, Joseph Banks, and the supporting production and clerical staff

(4) Special contributors and consultants—John J. Lottes and Gilbert Palmer, State University of New York, College of Education at Geneseo; Clyde Corcoran, Whittier College, Whittier, Calif.; James D. Gordon, University of Pennsylvania; Nelson W. Francis, Brown University (formerly at Franklin and Marshall College); H. M. MacNeille, Case Institute of Technology (formerly at Washington University, St. Louis, Mo.); Andrew V. Kozak, The Pennsylvania State University

Finally, deep appreciation is expressed to the Educational Media Branch of the Office of Education, Department of Health, Education, and Welfare for the financial support which made these research studies possible.

C. R. Carpenter

L. P. Greenhill

University Division of Instructional Services, The Pennsylvania State University

March 1963

SUMMARY

INTRODUCTION

The general purposes of this series of research studies were (1) to investigate the possibilities of programing entire courses in mathematics and English grammar, and (2) to compare different methods and media, with special emphasis on closed-circuit television, for presenting programed learning materials. In order to accomplish the objectives it was necessary to analyze especially the problems of pacing and individual and group conditions for the learning of programed courses.

The report deals specifically with the following subjects and problems:

1) A critical analysis of some assumptions of programed learning and implications for adapting and using the "new" media.
2) The development of programed materials for research purposes.
3) Experimental comparisons of external-pacing and individual-pacing for the study of programed materials using different methods and media for presentations.
4) Comparisons of different rates of pacing and of individual and group conditions of study.
5) Comparisons of programed televised presentations with (a) teaching machine presentations, and (b) instructor-taught classes.
6) Studies of the effects of the minimum group of two students studying programs in comparison with individual study.
7) Preliminary analyses of quantitative and verbal abilities of students relative to achievements using programed courses, and of a personality factor relative to these study conditions.

The content areas selected for programing were introductory college algebra and English grammar. Special programed courses were written, tested, revised, and adapted for presentation by teaching machines, programed textbooks, film strips, and closed-circuit television. The completed algebra course consisted of 30 units of instruction; the course on English grammar contained 14 units. Appropriate and reliable tests of learning were also constructed. In general, these included tests on each unit of the programed courses, mid-course tests, and end-of-course tests. Additional special purpose tests were used in the studies.

THE EXPERIMENTS

Six separate controlled experiments were conducted. Four of these used the algebra program, and two used the English grammar program. These experiments are briefly summarized below:

CONTEMPORARY ALGEBRA COURSE

Experiment I made comparisons between several methods of presenting the programed course—by teaching machines (self-paced), by programed textbooks (self-paced), and by film strips (externally-paced). Performance of students receiving the programed instruction was compared with that of students receiving direct instruction from an experienced teacher who covered the same topics by the lecture-discussion method.

Results showed no substantial differences in performance among the three programed treatments as measured by the sum of students' scores on the unit-tests and their scores on a comprehensive end-of-course test. However, on the unit-tests the three programed treatments combined produced significantly higher scores and lower variance than did the direct instruction by a good teacher. On the end-of-course test the means were not significantly different, but again the variance for the programed treatments was significantly lower than that for the direct instruction.

Experiment II investigated the effects of varying the pacing of externally-paced programed materials presented to groups of students by means of film strips. The initial times for each frame under the externally-paced conditions were based on the mean time taken for each frame by a sample group of students working on a self-paced schedule. In this experiment these base times were referred to as 100%, and groups were paced frame-by-frame through the programed course using times for each frame which corresponded to 80%, 90%, 100%, and 110% of these base times.

Results showed that within this range of variations in external-pacing there were no significant differences in student performances or in attitudes toward programed learning. Again, externally-paced and self-paced study conditions were found to be equally effective.

Experiment III involved a comparison between an externally-paced presentation of the programed algebra course using closed-circuit television and a self-paced presentation using teaching machines.

Results showed equivalent learning from both procedures as measured by summed unit-test scores and performance on mid-course and end-of-course tests.

Experiment IV investigated another aspect of the pacing and grouping problem—namely, the pairing of students for the study of programed materials. In this experiment half of the students were randomly assigned to pairs; the remainder of the class worked individually. It was hypothesized that interaction between the students in each pair might facilitate learning and provide more favorable motivation, including immediate social reinforcement, for programed instruction.

Results showed that the performance of the students in the two treatments was equivalent on the unit-tests, the mid-course, and the end-of-course tests. Further analysis of the results indicated that different learning abilities affected performance in the two different treatments. In the paired group verbal ability accounted for a substantially higher proportion of the variance than did quantitative ability, whereas in the group that studied individually the reverse was found to be true. Student attitudes toward programed instruction were neutral and not significantly different for the two groups.

ENGLISH GRAMMAR COURSE

Experiment V conducted in the English grammar course comprised two studies with two different populations. One of these involved a comparison between a closed-circuit television (externally-paced) presentation and a teaching machine (self-paced) presentation.

The second study compared the performance of students taught by means of the programed course presented over closed-circuit television with that of students who were taught the same subject matter by an experienced instructor who used the lecture-discussion method.

The results showed that in each study there were no substantial differences in students' performances as a result of the different treatments they received. However, in the second study student attitudes as measured by an attitude scale were significantly in favor of the instructor presentation.

Experiment VI was designed to investigate a further variation in conditions of learning which involved the pairing of students to study the teaching machine version of a programed course. A preliminary attempt was made to relate a personality characteristic of students to these conditions. In this experiment students were randomly assigned to two groups. They were paired on a personality variable of dominance-submissiveness. One group was paired so that each member of the pair was similar in this personality characteristic; i.e. the *most* dominant was paired with the *next* most dominant, and so forth throughout the group. In the second group students were paired who were dissimilar on this variable. In this group students were ranked from most dominant to least dominant, and in a group of 30, number 1 was then paired with number 16, and so forth.

Results showed no significant differences in the learning of the two groups on the several tests that were used. However, there was some evidence that differential abilities and characteristics of the learners contributed to performance in the two groups.

GENERAL CONCLUSIONS

1) It appears to be feasible to adapt and use those media which require externally-paced methods for presenting programed materials to *relatively* homogeneous groups of learners without adversely affecting their achievement.

2) Students of the populations sampled have a relatively wide tolerance for variations in pacing rate.

3) Closed-circuit television supplemented by printed materials can be used effectively for the presentation of programed instructional materials to groups of students. This opens up possibilities for the wide distribution of programed instruction and for the integration of programed sequences with lecture-demonstrations in televised instruction. The use of programed materials on television, supplemented by printed response forms, provides an excellent means for obtaining active participation on the part of all learners and for providing immediate reinforcement of their responses.

4) Results on the pairing of students to study programed materials equal those obtained with individual study methods. The effects of this pairing procedure on motivation and attitudes are not yet conclusive and require further investigation.

5) Sampling of two content areas, mathematics and English grammar, yielded essentially the same results on all experimentally studied variables.

6) In general, this research challenges the classical assumptions that the best method of using programed materials is that of self-pacing and individualized study. In other words, a learner's own speed may not be the optimum rate. The research shows some of the possibilities of using the "new" media and combinations of media with paced programs for instructing large numbers of students simultaneously in time-limited periods of instruction.

PART I

THE "NEW" MEDIA AND PROGRAMED LEARNING

by C. R. Carpenter

Introduction

The purpose of this chapter is to provide background and orientation to the experiments reported in the subsequent sections. Problems and possibilities are outlined for research and development on integrating and reciprocally adapting the "new" media and instructional programing methods, procedures, and techniques. More specifically, the question is raised of how the characteristics of a wide range of modern communication facilities can be adapted and used to apply the *essential* principles and to meet the requirements of programed learning. Furthermore, general research and development strategies are discussed to provide a background for understanding the research methods used in the reported investigations. The scope of this chapter is broader than the research studies being reported. Therefore, it is hoped that the problems raised and the perspective of this introductory chapter will suggest new lines and programs of research on the "new" media and systematically ordered learning materials.

General Assumptions on Programing Procedures

In time the original assumptions about operant conditioning formulated by Skinner (1) and extended to the control of verbal behavior and the learning of complex, "meaningful" intellectual proficiencies may be classified as *classical operant conditioning*. This designation may closely parallel the classical conditioned response principles experimentally established by Pavlov. In the same manner that it has been necessary to extend and modify classical Pavlovian conditioning to describe behavior more complex than reflex responses, it seems necessary to modify Skinnerian assumptions of the nature of operant conditioning to describe and control complex learning of scholastic and academic subjects involving the use of "new" media.

Although many modifications and restatements of the basic operant conditioning assumptions which initiated extensive research on programed learning and teaching machines have already been made, these assumptions nevertheless need to be liberally reexamined and restated once again as an orientation to research on "new" media and programed instruction.

Perhaps the principal, relevant, classical assumption is that the organization of programed stimulus materials requires that limited, cohesive "frames" of information be written and that these be ordered in sequences to reflect the inherent structure of the content and to provide high probabilities that the desired learning will occur.

The conceptual dimensions of the "frames" or steps should be accurately gauged to interact favorably with the learners' knowledge and abilities. Some of these dimensions are: the amount of information or scope of the problem put in a "frame"; the kinds of information and its structure; the amount of partially known and new information; the level of its difficulty for students, and the responses demanded by each step of the sequences of stimulus materials. Although the art of programing specific kinds of content may result in many different frame patterns, these conceptual dimensions must be estimated (and usually modified subsequently as the result of pre-tests of student reactions) in order to produce effective programs.

It is implicitly assumed that the organized sequences of content should progress at an appropriate *rate of development* to insure expanding comprehension and increasingly successful learning performances by students. This rate of development of a subject is a resultant of organizing the materials, of managing the concept density, and of controlling the rate of progression from familiar to unfamiliar content, or from levels of less difficulty to levels of gradually greater complexity and difficulty.

The *pacing rate* is different from, but interacts with, the rate of development. The latter involves the manner in which the content is composed in frames and the sequential integrity of these frames, which constitute the units or periods of study. Pacing is definable as the regulation of time spent on various phases of the stimulus-response-reinforcement cycle.

These assumptions lead to another which has become a programing doctrine; namely, that the pacing rate should be individually regulated by the student himself. In other words, it is assumed that the student should be permitted to proceed through programed materials at his own speed. This requirement for individualized pacing, however, does *not* include control by the learner of the *rate of development* of the subject matter, although the two sets of rates certainly interact. For example, a very gradual rate of development may permit a more rapid pacing; and conversely, a rapid rate of development may result in the student's slowing down his pacing rate. Thus, the student, under classical assumptions, controls the pacing rate, but the rate of development of the content in programing is a function of the program writing, based, of course, on the results of repeated tests of performances.

The concept of pacing can be considered in several categories:

(a) Self-pacing with unlimited time. Under this condition the students work individually at their own speed and take as much time as they need to complete each unit of the program.

(b) Self-pacing, time-limited. Under this condition the students work individually, but with a time limit for the completion of each unit of the program. The time allowed is sufficient for all learners to complete the program if they keep working, but the rate is affected by the set limits of time.

(c) External pacing. Under this condition the students work through the program as a group, sufficient time being allowed on each frame for the students to read it, respond to the question or problem, and check their responses. The times are generally determined empirically.

Silverman and Alter (2) have conducted a study which investigated the problem of pacing in a preliminary way. They arranged students in three groups—paced, long paced, and unpaced. No significant differences in performance were found. However, the results of this study cannot be considered as conclusive because the program was a short one of 87 frames, there were only 15 students in each group, and the criterion measure contained only 22 items. The study does, however, challenge the assumption of the superiority of self- or individual-pacing as opposed to external-pacing.

In the studies to be reported here comparisons will be made between the performances of students working as individuals under self-paced, time-limited conditions using teaching machines or programed textbooks, and those receiving externally-paced group presentations by means of film strips or television. These studies involve a number of units of programed instruction (in some cases entire courses) in two subject matter fields.

"Branching programs" as developed by Crowder (3) provide another kind of control, within limits, in that different patterns of content and different tracks and sequences of materials are followed as indicated by the students' performances, errors, successes, and preferences.

As linear analysis techniques are applied to programing operations, the dichotomy between linear and branching types of programs is disappearing. Furthermore, adequate results and records of student performances permit the designing of program loops, cycles, regressions, advances, shifts of levels of difficulty, and other patterns which meet the learning needs of different students. Therefore, efficient *pre-programing* of many kinds of formats becomes possible. With increased information on student performances, furthermore, the precision of program regulation is increased, and more of the control is shifted from the student back to program writers and producers.

In this connection, a variation of linear programing which invites attention is the repetition of difficult learning cycles for those who require practice in order to meet specified performance standards. Using varied repetition (as is required, for instance, in language learning), *pre-programed* repetition and reviews of critical stimulus-response-reinforcement cycles may be provided in materials prepared for use by the "new" media.

Another classical programing assumption relevant to this research, but not peculiar to this method except in degree of emphasis, is that arranged content and conditions for learning require continuous or consecutive actions and reactions by students. In Skinnerian terms, behavior must be elicited overtly. It is assumed, that there should be *persistent*, *intense*, and *appropriate* student reactions to the stimulus materials. The practical problem then is how to arrange conditions to insure that *persistent*, *intense*, and *appropriate* reactions will occur. To this end, each frame requires a response. The frame sharply limits the kinds and ranges of response alternatives. Two among many possible kinds are "constructed" and "selected" answers. The structure of the frame includes previously established response "sets" or instructions, cueing suggestions, and positioning of items of information which establish a high probability that the desired response will be made.

Classical operant conditioning assumes that the responses must be overt—not covert—in order to have "expressed behavior" which can be observed, recorded, and used to regulate subsequent behavior. As with the other original assumptions, that of the necessity of overt responses is being challenged by research (4, 5).

Closely related is the assumption that correct responses should greatly predominate over errors. The extreme position is that there should be no errors in response to perfectly programed materials. The practicability of meeting this requirement through the use of small easy steps, dense cueing, and slow *rates of development* is also being challenged. The problem can be stated more precisely as that of determining and using the optimal ratio of correct responses to errors for specified kinds of content and the required terminal performances of students.

Knowing whether, or to what degree, a response or solution is correct seems to be the important contingency for learning. The learner's knowledge of the results of his efforts to learn appears to be the essential condition, and how and in what form this information is provided to the student is only secondary. It can be accomplished in many different ways in the art of programing content, or indeed in the management of other kinds of instruction (6).

This discussion leads to another assumption; namely, that the "feedback" or knowledge of the logically inherent or ascribed *value of a response* be given the student *immediately* after he reacts. The technical term is "immediate reinforcement" of correct responses. Knowledge of the value of a response is two-edged: correct responses are *reinforced* and errors are *extinguished.* By controlling and applying these positive and negative effects, in Skinnerian terms, specified behaviors are "shaped," or in common language, "learned."

The timing of reinforcement and extinguishing cues is known to be important, and the provision for supplying information about the value of learning reactions *while these reactions are perseverating* appears to be one of the main advances made by programing methods. The timing procedures used in programed instruction are very different from the delayed, often ambiguous knowledge of results of student performances which characterizes conventional methods of instruction.

The assumptions of classical linear programing reviewed above are not the only ones that could be listed, but it is believed that they are the principal ones relevant to the research being reported. These assumptions have special implications for the problems of integrating and reciprocally adapting programing methods and some of the "new" media, especially closed-circuit television.

Implications of the General Assumptions for Functions of Teaching Machines

The assumptions which have been reviewed can be translated into functional requirements and performance specifications for teaching machines and other kinds of instructional instrumentation. The development of the technology of instruction and learning has proceeded apparently on the assumption that especially designed machines are necessary to satisfy the functional requirements for storing programed materials, for retrieving these frame by frame, for providing appropriate responses, and for giving correctly timed reinforcement or response evaluative cues. Machines and books have been especially constructed for the branching types of programs and for many kinds of special discrimination and performance skills learning. Machines have been built for presenting ordered printed programs on cards of varied sizes and paper in sheets, folds, and rolls. Machines have been designed to use film strips, slides, transparencies, graphics, and, to a limited extent, audio tape (7).

Programing methods have already (1962) led to the production of new equipment and arrangements for managing programed materials by about 140 individuals, agencies, institutions, businesses, industries, and publishers, according to Finn and Perrin (8).

Publishers, with few exceptions, have argued that the *essential* functional requirements for programed instruction can be met by appropriately designed books, and they have resisted investments in the new kinds of teaching machines. Finn and Perrin define programed books as "simulated" teaching machines. Indeed, the question of what essential functional requirements for programed instruction cannot be implemented by especially designed programed books is central. And, more generally, what is the relative effectiveness of different ways of presenting such materials by many different kinds of instructional instrumentation and media?

The effort to adapt textbooks to the requirements of programed learning is a noteworthy exception to the general trend toward building *new* teaching machines rather than stressing the development of existing apparatus, including the available electronic media. Generally, also, many of the "new" media, with their varied characteristics and potentials, have been ignored. The promising results, for example, of programed sound film loops (9) and film tests (10), and arrangements for learner participation (11, 12, 13) appear not to have been adequately evaluated, even as primitive prototypes, by teaching machine enthusiasts.

Neglect of "New" Media for Programed Learning

The factors that have contributed to the neglect, or at least lack of emphasis, on developmental research on media for use in programed learning result from the classical programing assumptions and doctrines. First, there is the assumption that the procedures must be individualized. Narrowly conceived, this assumption specifically implies that the student must control his own *pacing rate*. Sometimes a derived assumption—that by permitting time spent on programs to vary from individual to individual, good programs with almost perfect responses and very low error rates will tend to equalize the achievement of students of different abilities and backgrounds in the subject matter—is made. If permitting *time* to vary for individuals can really equalize student performances with good programs, then the individualized control of pacing rates is of fundamental importance. The validity of these assumptions remains to be proven. Second, the "new" media are usually stereotyped as *mass* media, appropriate only for group instruction, and group instruction is obviously antithetical, it is reasoned, to individualized instruction.

Three points are pertinent: first, that all learning is primarily personal and individual; second, that even though students may be taught in groups of various structures and sizes, learning remains personal or individual; third, that the programing doctrine of individualized learning is consistent with some educational philosophy, especially for elementary school levels, and therefore has had an appeal that exceeds the supporting evidence.

With few exceptions the programed learning movement was pioneered by men who did not work in the media field. Neither the film programs at Yale, which conducted important research on "participation" of learners with film and film strip materials (14), nor the Instructional Film Research Program at Penn State, which did research on "rates of development" (15), what would now be called programed sound film loops for language instruction (16), classroom communication "feedback" systems (17), knowledge of results (18), organization of content (19), and sound film test-teach prototypes in performance training (20), became points of departure for the present programing movement. Thus, programing originated mainly outside of the context of audio-visual and media research and development.

The lack of research data on the comparative effectiveness of different ways of using programed materials has contributed to the neglect of "new" media in programed learning. When it is *not* known whether, with the same content, alternative ways of presenting the materials result in learning differences, then assumptions and doctrines tend to exclude non-individualized methods from evaluation. Only recently have controlled comparative effectiveness studies begun, and the research being reported here will add needed evidence in this area.

Functional Requirements of Devices for Presentation of Programed Learning

Finn and Perrin (21) have proposed systematic definitions of teaching machines and have established useful classifications: audio-visual machines which have no integral response mechanism; those that provide for written responses to printed or projected media; those having key responses to written and projected media; the class of machines which have keyboards and require constructed answers; and the classroom communicator systems.

The following is a listing of the *general* functions which may be provided for in the design of machines, simulated machines, and systems of instructional instrumentation:

1) Provide for *storage* of organized programed information.
2) Provide for sequential, patterned, or random access to stimulus materials or "frames."
3) Display stimulus materials legibly and clearly to one or more of the sensory-perceptual modes of the learner.
4) Be flexible and adaptable to permit use of a range of different content channels, signs, and symbols.
5) Provide for appropriate student responses relative to performance requirements and criteria.
6) Present immediately to the learners, or during a critical period, adequate means of assessing the value of their responses ("Feedback," reinforcement, extinction, etc.).
7) Regulate appropriately the rate and repetition of learning cycles (or steps) for efficient learning.
8) Record data on the essential phases of the stimulus - response - reinforcement cycle.

This *general* set of functions is believed to describe the characteristics of the majority of teaching machines and devices. Many of them do not provide for all of the general functions listed. Some are so constructed as to emphasize certain functions (e.g., phasing-in of reinforcement or the use of print characters), but neglect other functions (e.g., selective data recording or adaptability for a range of different kinds of content and symbols).

If it can be agreed that the functions listed above implement the essential assumptions given earlier and that the functions generally describe most (but not all) teaching machines, then this Introduction can logically proceed to the next step. This step is to examine a range of the actual and potential characteristics of the "new" media and to inquire about the possibilities of adapting and developing them, *either singly or in combination*, for implementing the assumptions and serving the general functions of programed learning.

Characteristics of "New" Media Relative to Programed Learning

Existing communications media as a group have very broad ranges of capabilities for storing organized information, for patterned retrieval, and for displaying stimulus materials in one or more perceptual modes. These general functions are also possible characteristics of teaching machines and other arrangements for administering programed materials.

Static visual stimulus materials such as still pictures, graphs, and print can be stored, retrieved, and presented in varied patterns and sequences by slides, film strips, and transparencies. The stimulus materials can be in black and white or color of high technical quality, and the sizes of displays can be varied appropriately for clear viewing by a single individual or groups including large audiences.

Dynamic visual stimulus materials can be stored, retrieved, and displayed by motion pictures, television, and video tapes. Here, too, color may be used, and static visuals and print can be incorporated.

Audio stimulus materials can be presented by disc and tape recorders and reproducers, by radio, and by special equipment assemblies, as in language laboratories.

The greatest possibilities for developing and adapting the "new" media for use with programed learning become evident when combinations of capabilities of several media are considered, as, for example, when the static and dynamic video media are considered in combination with print either carried by film or video tape or on paper, or when sound recordings are added to visual materials.

Motion pictures and television (including materials that supplement films and video tape recordings) not only provide broad bands or channels for audio and visual stimulus materials with great ranges of communicative characteristics, but the distributive possibilities, especially in the case of television, are enormously extensive. These are *general* media with which many other specific audio-visual media are compatible, and in addition they can be used with printed materials.

All of the media, either in the production of instructional materials or in their presentation-display functions, have possibilities for controlled *rates of development* of the content and for pre-programing the *pacing* of presentations.

The "new" general media, especially motion pictures and television, are also adaptable to a wide range of content, ranging from the very abstract and symbolic to the accurate reflection or simulation of concrete objects and situations.

There are other general functional requirements which are not met by the "new" media as they have been used conventionally. However, it is proposed that these functions can be provided for in three major ways: first, by creating new types and formats of materials for mediation by existing audio-visual equipment; second, by developing patterns of integrated programs for use by combinations of media apparatus, e.g. film strips, sound tapes, and print; and third, by designing auxiliary equipment systems for serving essential functions that are not served by the existing media apparatus.

The general functions of providing for appropriate responses, presenting information for assessing the value of responses (reinforcement and "feedback"), regulating the rate and repetitions of learning cycles as required by students, and obtaining essential permanent records of certain phases of the learning cycle and student responses may need to be provided by one or all of the three approaches suggested above.

Finally, there are some assumptions basic to programed instruction and the corresponding derived functional or implementing requirements, which the "new" mass media cannot fully satisfy within the limits of feasibility or practicality. Several of these functions are individualized learning, including the control or regulation of pacing rates of the stimulus materials or learning cycles concurrently by the individual learner as he studies. The question arises of how essential these assumptions and functions are and what *tolerance ranges* for variations exist from the classical views of programed learning without causing unacceptable decrements in the effectiveness of learning. The important question is how variations of methods *within tolerance limits* of students will affect the efficiency of learning.

In general, these are questions which the research projects being reported seek to answer. The main emphasis was on explorations for reciprocally adapting programed learning and closed-circuit television. The discussion of the "new" media in a general perspective is brought into focus on programed learning and closed-circuit television when it is proposed that television be used as the mediating system or teaching machine.

Adapting Television and Programed Courses

Those who are familiar with the content and program productions now used in instructional and educational television are aware of the needs for inventiveness, knowledge about conditions of learning, and creative imagination in developing more effective, efficient formats and patterns of instruction for television presentation. The problem is often stated as one of making the best use of the full potentials of the television medium, which means the production of instructional materials which create the highest probabilities that desired kinds and levels of learning will occur. It has long been known that neither the reflection of conventional classroom methods and procedures nor the employment of entertainment-commercial film and television formats are answers to the problem. New and better kinds of programs are needed.

The idea of preparing and presenting materials which are adaptations of programed instructional methods and materials is a development very different from current television production methods. The attempts to introduce programed learning materials into instructional television courses and operations may at least serve as a kind of "shock therapy" for television instructors and producers. Similar needs exist for instructional films and radio.

More importantly, the presentation of programed materials over television offers a possible means of obtaining frequent, active, overt responses from all learners, and of giving them immediate knowledge of results—something which is rarely achieved in large-group instruction and the absence of which is frequently cited as a "disadvantage" of television. Also, the use of television for the presentation of programed materials, if successful, would not only overcome the logistical problems of providing programed materials for large numbers of students, but the flexibility of television should make possible the integration of a wide range of instructional resources into a program: e.g., live televised instruction, films, pictures, demonstrations, etc., as well as programed sequences.

In this connection Gropper and Lumsdaine (22-27), in cooperation with the Metropolitan Pittsburgh Educational Television stations WQED-WQEX, have recently conducted a series of studies which involved the adaptation of programed materials to televised instruction. Some of these studies have compared the effectiveness of televised lessons containing programed sequences with conventional televised lessons. Others have investigated various aspects of student response to programed TV instruction.

Since television instruction cannot be feasibly provided for single individuals, a second crucial problem is that of determining whether and with what degrees of relative effectiveness *groups* of students can be taught over television using programed materials, as compared with individualized methods and procedures. To collect evidence on this problem, several research studies have been conducted and will be reported in subsequent chapters.

Although the rate of development is determined by program writers in the same manner for individualized or for group uses of materials, the pacing rate must be *preprogramed* for groups or classes of students. The problem is a procedural one of determining the ranges and tolerance limits of pacing rates which will make it possible for large groups of students to engage themselves *intensively, persistently,* and *appropriately* in programed learning.

The classical programing assumption that the pacing rate a learner uses *is the best rate for his learning* is challenged. The rates at which students read are neither fixed nor optimum in most cases. Reading-comprehension rates can be improved by training. Habitual patterns of study and learning are also subject to improvement by training. It is reasoned that controlled pacing using gradual increments or *lead rates* which students can accept, could significantly increase the rates actually preferred by students.

These considerations justify the research effort emphasized in the projects being reported for developing methods for *pre-programing pacing rates* for groups and the making of comparisons among different patterns of rates. It is necessary to solve this problem before programed learning can be mediated feasibly by television. However, the evidence on pacing rates has broader implications, both theoretical and practical, than just for television instruction.

Another crucial problem is that of determining what essential programing assumptions or functions can be served by special productions and television systems and what functions need to be mediated in other ways. Here the multi-media approach is indicated. This approach requires that combinations of media be selected and that the systems approach be used in determining what different media in what combination are needed for the task.

One analysis of programing functions yields four categories related to adapting ordered materials for use over television: (1) display and presentation of frames or steps of information; (2) student responses; (3) reinforcement; (4) the recording of student response data.

Familiarity with television operations as well as feasibility tests shows that functions (1) and (3)—display-presentations and reinforcement—using either the video or audio channels or both can be very effectively served by using television. Functions (2) and (4), providing for student responses and their recording, require paper-pencil forms synchronized with the program content, or special "classroom communication" equipment.

Non-programed introductions or summaries and the descriptions of unit and subunit learning requirements can also be effectively presented over television to students, interspersed with programed materials. Tests and examination questions and problems can be clearly presented over TV, and the students' answers or solutions can be recorded on answer sheets or printed forms.

In brief, television is a general communications medium for which many kinds of stimulus materials are appropriate, ranging from very abstract symbols (including print) to real objects, persons, and situations. Speech, music, and other sounds can be mediated. Many other specialized materials such as demonstration equipment, display cards, slides, film strips, motion pictures, film loops and programed sequences can be used in conjunction with television. In the future, programed course materials can be recorded on video tape and distributed by television either in closed-circuit or broadcast operations. These possibilities provide wide latitude for the creative work of program writers and planners and make it possible to prepare different kinds of content for use over television.

In summary, the two most crucial issues involved in reciprocally adapting television and programed learning are whether the pacing rate problem can be solved; and if so, whether groups of students who have similar characteristics can be taught as effectively as those receiving individualized programed instruction. This report will provide evidence on these issues.

Research Strategy

There are general research strategy considerations relative to the studies being reported which will help the reader understand the problems, methods, and results.

A small research laboratory was established to conduct this project. This laboratory was committed almost entirely to course development work. The general strategy was to define an academic area which presented instructional problems or needed improvement and to cycle this area into the research laboratory for the concentrated effort necessary to develop it to the highest possible standards. The idea then was to field-test the courses and reintroduce them into the regular academic curriculums.

This plan of course development, it is suggested, provides the necessary conditions for course content and methods improvement. It is rarely possible to provide the necessary conditions for course development while the course is being given.

The sample of content to be programed received early attention. The research literature reports many studies done on limited numbers of frames (28). It was determined early to use entire courses of instruction as the *content sample* and to avoid the possible undesirable consequences of small samples of content.

Mathematics and English course content was selected. Efforts were made to insure that the content of the programed courses prepared for these studies met high professional standards. Committees of professional associations in mathematics and English and the academic departments concerned were asked for help and invited to designate consultants. The attempts to insure high quality of programing and production techniques were more uncertain in results because of the absence of proven standards and examples of programed materials in 1960-1961. In particular the print and symbol requirements for both mathematics and English limited efforts to exploit the full video potentials of television and auxiliary media.

The strategy of conducting research with students and courses in regular academic curriculums was employed. This type of contextual research, although difficult to accomplish, insures both a pattern of motivation desirable in research and also the availability of student subjects in numbers necessary for statistical treatment of data. It also possibly increases the utility and generality of the products of research and development.

An interdisciplinary team approach was used in the research being reported. Content people wrote the programed courses with training and assistance from others. Support was provided in planning, experimental design, tests and measurements, production services, and in the conduct of the course instruction after the materials were prepared and duplicated. The whole job adds up to a massive effort over a period of two and one-half years by a group of highly motivated people with different competencies.

Summary

The research studies to be reported here were designed to investigate the possibilities of presenting programed courses to groups of students via closed-circuit television and to compare this procedure with other methods of presenting programed courses.

In particular, the following problem areas were investigated:

1) group presentation of programed courses vs. individual presentation

2) the varied pacing of group presentations of programed instruction

3) comparison of televised and teaching machine presentations of programed courses with face-to-face instruction

4) the effects of pairing students to work on programed course materials (a variation of the grouping problem with added possibilities of interaction)

The subject areas selected for programing were introductory college algebra and English grammar. Special programs were written, tested, revised, and adapted for presentation by teaching machines and programed textbooks (self-paced), and by film strips and closed-circuit television (externally-paced).

The experiments in the programed mathematics course will be described first, followed by a discussion of the experiments in programed English grammar.

References

(1) Skinner, B. F., *The Behavior of Organisms*, Appleton-Century-Crofts, New York, 1938.

(2) Silverman, Robert E., and Alter, Millicent, *Response Mode, Pacing and Motivational Effects in Teaching Machines*, Technical Report NAVTRA DEVCEN 507-3, U.S. Naval Training Device Center, Port Washington, New York, 1961.

(3) Crowder, Norman A., "Automatic Tutoring by Means of Intrinsic Programing." In *Automatic Teaching, The State of the Art* (ed. Galanter), Wiley, New York, 1959.

(4) Ashbaugh, William H., *The Effect upon Achievement of Written Responses to Programed Learning Material for Students of Differing Academic Ability.* Unpublished doctoral dissertation, The Pennsylvania State University, 1962.

(5) Silverman, Robert E., and Alter, Millicent, *op. cit.* (Ref. 2).

(6) Moore, J. William, and Smith, Wendell I., *The Role of Knowledge of Results in Programed Instruction.* Mimeo, Bucknell University, 1962.

(7) Morton, F. Rand, *The Language Laboratory as a Teaching Machine*, The Language Laboratory, The University of Michigan, 1961.

(8) Finn, James D., and Perrin, Donald G., *Teaching Machines and Programed Learning, 1962, A Survey of the Industry.* Occasional Paper No. 3, Technological Development Project, National Education Association, Washington, D. C., 1962.

(9) Kale, S. V., and Grosslight, J. H., *Exploratory Studies in the Use of Pictures and Sound for Teaching Foreign Language Vocabulary*, Technical Report SDC 269-7-53, The Instructional Film Research Program, The Pennsylvania State University, 1955.

(10) Carpenter, C. R., Greenhill, L. P., Hittinger, W. F., McCoy, E. P., McIntyre, C. J., Murnin, J. A., and Watkins, R. W., "The Development of a Sound Motion Picture Proficiency Test," *Personnel Psychology*, Vol. 7, No. 4, 1954.

(11) Roshal, S. M., *Effects of Learner Representation in Film-Mediated Perceptual-Motor Learning*, Technical Report SDC 269-7-5, The Instructional Film Research Program, The Pennsylvania State University, 1949.

(12) Ash, P., and Jaspen, N., *The Effects and Interactions of Rate of Development, Repetition, Participation and Room Illumination on Learning from a Rear-Projected Film*, Technical Report SDC 269-7-39, The Instructional Film Research Program, The Pennsylvania State University, 1953.

(13) May, Mark A., and Lumsdaine, Arthur A., *Learning from Films*, Yale University Press, New Haven, 1958.

(14) May, Mark A., and Lumsdaine, Arthur A., *op. cit.* (Ref. 13).

(15) Jaspen, N., *Effects on Training of Experimental Film Variables. Study I: Verbalization, Rate of Development, Nomenclature, Showing of Errors, "How-it-Works," Repetition*, Technical Report SDC 269-7-17, The Instructional Film Research Program, The Pennsylvania State University, 1950.

(16) Kale, S. V., and Grosslight, J. H., *op. cit.* (Ref. 9).

(17) Carpenter, C. R., Eggleton, R. C., John, F. T., and Cannon, J. B., *The Classroom Communicator*, Technical Report SDC 269-7-14, The Instructional Film Research Program, The Pennsylvania State University, 1950.

(18) Hirsch, R. S., *The Effects of Knowledge of Test Results on the Learning of Meaningful Material*, Technical Report SDC 269-7-30, The Instructional Film Research Program, The Pennsylvania State University, 1952.

(19) Northrup, D. S., *Effects on Learning of Organizational Outlines in Instructional Films*, Technical Report SDC 269-7-33, The Instructional Film Research Program, The Pennsylvania State University, 1952.

(20) Carpenter, C. R., Greenhill, L. P., *et al.*, *op. cit.* (Ref. 10).

(21) Finn, James D., and Perrin, Donald G., *op. cit.* (Ref. 8).

(22) Gropper, G. L., and Lumsdaine, A. A., *An Experimental Comparison of a Conventional TV Lesson with a Programed TV Lesson Requiring Active Student Response*, Studies in Televised Instruction, Report #2, American Institute for Research, 1961.

(23) Gropper, G. L., and Lumsdaine, A. A., *An Experimental Evaluation of the Contribution of Sequencing, Pre-testing and Active Student Response to the Effectiveness of Programed TV Instruction*, Studies in Televised Instruction, Report #3, American Institute for Research, 1961.

(24) Gropper, G. L., Lumsdaine, A. A., *An Investigation of the Role of Selected Variables in Programed TV Instruction*, Studies in Televised Instruction, Report #4, American Institute for Research, 1961.

(25) Gropper, G. L., and Lumsdaine, A. A., *Issues in Programing Instructional Materials for Televised Presentation*, Studies in Televised Instruction, Report #5, American Institute for Research, 1961.

(26) Gropper, G. L., and Lumsdaine, A. A., *The Use of Student Response to Improve Televised Instruction—An Overview*, Studies in Televised Instruction, Report #7, American Institute for Research, 1961.

(27) Gropper, G. L., *The Role of Visuals in Verbal Learning, An Interim Report*, American Institute for Research, 1962.

(28) Lumsdaine, A. A., and Glaser, R. (eds.), *Teaching Machines* and *Programed Learning—A Source Book*, National Education Association, Washington, D. C., 1960.

PART II

RESEARCH ON METHODS AND MEDIA FOR PRESENTING PROGRAMED MATHEMATICS

The Development of the Mathematics Program

The scope, content, and sequence of the programed units to be used in the mathematics course were determined early in the fall of 1960 after consultation with a group of nationally known mathematicians who were serving on the National Academy of Science's Committee on Mathematical Films and Television. It was decided that the content should be "modern" in character; accordingly, the recommendations of the Commission on Mathematics and the published materials of the School Mathematics Study Group had considerable influence on the selection of the topics for the course. In addition, the scope and content of the course units were selected so as to be acceptable to the Mathematics Department of The Pennsylvania State University for use in Mathematics 2 (College Algebra). The programed course that was finally developed, including the units used in the experiments, consisted of 30 units entitled "Contemporary Algebra." The titles of the various units appear in Table 1. The number of frames in the various units ranges from 82 to 156, the total number of frames being 3,770. The program is of the linear, constructed-response variety.[1]

The development of the course consisted of three distinct but overlapping phases. A description of these phases of work follows:

PHASE I: (October, 1960-April, 1961)

After the first writing of each of the programed units, they were administered individually to four students under the direct observation of one of the exerimenters.[2] The observer made a record of each frame for which there was an incorrect response. He discussed the content of each of the units with each student in an effort to ascertain the extent to which the student had learned the concepts that were presented. The results of these intensive observations were studied by the writers and used as the basis for making the first revision of the units.

TABLE 1

TITLES OF PROGRAMED UNITS

Unit	*Title*
1.	Numbers
2.	Sets
3.	Sets and Variables
4.	The Language of Algebra (Introduction)
5.	Operations—Part I
6.	Operations—Part II
7.	Arithmetic of Real Numbers—Part I
8.	Arithmetic of Real Numbers—Part II
9.	Arithmetic of Real Numbers—Part III
10.	Linear Equations and Inequalities in One Variable—Part I
11.	Linear Equations and Inequalities in One Variable—Part II
12.	Linear Equations and Inequalities in One Variable—Part III
13.	Linear Equations and Inequalities in Two Variables—Part I
14.	Linear Equations and Inequalities in Two Variables—Part II
15.	Systems of Linear Equations—Part I
16.	Systems of Linear Equations—Part II
17.	Systems of Linear Inequalities and Word Problems
18.	Polynomials and Factoring—Part I
19.	Polynomials and Factoring—Part II
20.	Quadratic Equations and Inequalities
21.	Rational Expressions
22.	Rational Expressions and Equations
23.	Exponents
24.	Exponents and Radicals—Part I
25.	Exponents and Radicals—Part II
26.	Radicals and Equations
27.	The Quadratic Formula
28.	Complex Numbers
29.	Functions and Relations—Part I
30.	Functions and Relations—Part II

PHASE II: (January, 1961-June, 1961)

The revised program was tested with a class of approximately 30 pre-service elementary school teachers who were enrolled in a Mathematics 200 course during the Spring Semester, 1961, at The Pennsylvania State University.[1] These students completed one unit each week in place of one of three regularly scheduled lectures. Results on the error rates, both by frame and by unit, were collected, as well as performance information on tests which were prepared to accompany each of the various units.[2] These data were analyzed, and the information was used as a basis for a second revision of the first 15 units of the program.

---

[1] The program was written by Ralph T. Heimer and Frank Kocher.

[2] John J. Lottes, now of the State University of New York, College of Education at Geneseo.

PHASE III: (June, 1961-September, 1961)

The final testing of the entire program took place during the summer of 1961 with a group of 14 students, most of whom were high school seniors in the College Area School System, State College, Pennsylvania. Additional data, similar to the data of Phase II, were assembled, and an electronic operations recorder, which provided a response station for each student, was used to record the time required by each individual to complete each frame cycle, as well as the total time required to complete each unit. These data were studied and the final revision then undertaken. The timing data were used both in an effort to improve the effectiveness of the programed material and to adjust the content in the units so that each could be completed in 65 minutes or less.

A member of the Mathematics Department of The Pennsylvania State University[3] read the entire program for content accuracy and made editorial changes during the several stages of its development.

Except for minor alterations, the conclusion of Phase III marked the completion of the development of the programed course material.[4]

---

[1] Under the supervision of Professor A. V. Kozak, who made many helpful suggestions.

[2] The tests were prepared by John J. Lottes.

[3] Albert Erskine.

[4] The program is now published by Holt, Rinehart and Winston, Inc. of New York City under the title *Contemporary Algebra.*

DEVELOPMENT OF ANCILLARY MATERIALS

To supplement the programed units, study guides were prepared—one to accompany each of the 30 units.[1] These study guides were written during the summer of 1961 and were first used during the third phase of the development of the program. Following this first use, the study guides were revised and used consistently to supplement the units in all research projects.

Each study guide had six main sections:

I. **Vocabulary:** this section listed vocabulary words that were introduced in the corresponding unit and defined each word.

II. **Symbols:** this section gave the new symbols that were introduced in the corresponding unit and explained each symbol.

III. **Some Comments and Important Results:** this section listed and explained important concepts that were discussed in the corresponding unit.

IV. **Additional Discussion:** this section was a discussion which gave additional explanation of the important concepts presented in the corresponding programed unit.

V. **Sample Problems:** this section consisted of carefully selected problems, worked out in detail, which illustrated the important ideas and techniques discussed in the corresponding unit.

VI. **Problems:** this section consisted of problems to be used for assigned individual work out of class.

DEVELOPMENT OF EXAMINATIONS

A large pool of examination questions (one set for each of the programed units) was written and tested during the third phase of the development of the course. All of the questions were of the multiple-choice variety.[1]

---

[1] Most of the study guides were prepared by Clyde Corcoran and Gilbert Palmer.

After this preliminary testing of the items, *unit-tests* (which require a mean working time of approximately 10 minutes) were constructed to cover each unit of the program. An attempt was made to balance the items in each unit-test with respect to content, difficulty, and discriminative power. Two *comprehensive examinations* were also constructed using the item pool, one covering the content of the first 15 units and the other the content of the entire course.

Each of the comprehensive examinations contained 60 questions; the number of items in the unit-tests ranged from 11 to 19. The reliabilities of the examinations were computed and are reported in Table 2, below.

The philosophy underlying item construction was that rote memory should be minimized and testing for understanding maximized. Emphasis also was placed on fine discrimination in both vocabulary and knowledge of concepts in "application" type items. Deliberate attempts were made to present many variations of applications.

Many items were designed to measure depth of understanding. This was accomplished by presenting problems which required a higher level of abstraction than the program problems demanded and also by requiring the student to perceive limitations imposed upon the application of principles in some circumstances.

TABLE 2

RELIABILITY COEFFICIENTS OF THE CRITERION MEASURES

Criterion Test	Reliability Coefficient[2]
Summed Unit-Tests	.94
15-Unit Comprehensive Examination	.86
Final Comprehensive Examination	.81

---

[1] The tests and examinations were constructed by John J. Lottes.

[2] Reliabilities were computed by Kuder-Richardson Formula 20.

EXPERIMENT I: Comparison of External-Pacing and Self-Pacing of Programed Instruction in Mathematics Using Different Methods of Presentation

by R. T. Heimer

PURPOSE

The primary purpose of this experiment was to determine whether presentation of programed materials to groups of students under externally-paced conditions is an effective procedure.[1] This experiment was conducted first in the sequence of experiments to be reported since the adaptation of programed instruction for television presentation challenges the classical assumption of programed instruction which requires self-pacing by each learner. Because the problem of pacing is critical, and also because Experiments II, III, and IV were similar in methods, materials, and analysis, the description and discussion of this first experiment will be presented in detail.

A secondary, but necessary, purpose of Experiment I was to assess the overall effectiveness of the programed materials. Every effort was made to write a high-quality program, but a more formal evaluation was required.

STATEMENT OF THE PROBLEM

The problem was divided into two parts:

Part I was intended to assess the effectiveness of the programed materials. Stated in more operational terms, this sub-problem was to compare experimentally three different methods of presenting the programed algebra course with a conventional presentation of the same content, using unit-tests and comprehensive achievement measures.

---

[1] See also, R. T. Heimer, *The Preparation of A Program in Contemporary Algebra and a Study of Its Effectiveness for Group Instruction Under Paced Conditions.* Unpublished doctoral dissertation, The Pennsylvania State University, 1962.

The following diagram illustrates the experimental comparisons to be made:

TABLE 3

EXPERIMENTAL DESIGN
FOR ASSESSING PROGRAM EFFECTIVENESS

Programed Presentations	vs.	Conventional Presentation
Self-paced programed books		Lecture-Discussions
Self-paced teaching machine mediated program		(covered same topics as programs)
Externally-paced programed film strips		

Part II was designed to test the hypothesis that self-pacing is not a critical factor in the effectiveness of programed materials. In operational terms, the problem was to compare the relative effectiveness of three methods of presenting identical programed materials:

1) programed book presentation (self-paced)
2) teaching machine presentation (self-paced)
3) group presentation by film strips (externally-paced)

It should be noted that treatments (1) and (2) above represent standard methods for using programed materials. Treatment (3), however, is the presentation of primary concern in this research. Unlike the usual procedure for use of programed materials, the film strips are here presented to *groups* of students under externally-paced conditions.

METHOD

Subjects

The population which was employed in this research consisted of all students who enrolled for a course in Mathematics 2 which met the first period[1] on Tuesday, Thursday, and Saturday mornings during the Fall Term, 1961, at The Pennsylvania State University. Initially there were 113 subjects involved, most of whom were freshmen with a major in a non-science area (Agriculture, Business Administration, Education, etc.).

Materials

The instructional materials used in this study consisted of the first 15 units (and the accompanying study guides) of the 30-unit algebra program previously described.

The accompanying unit-tests and mid-term comprehensive examinations were used as criterion measures. The Kuder-Richardson estimate of reliability for each test is given in Table 3.

Each unit-test was administered following the completion of the unit. In the self-paced treatments each student was given the unit-test as soon as he completed the unit. Under externally-paced conditions the unit-test was administered as soon as the group presentation of the unit was completed. In the conventional treatment the unit-tests were given during the last 10 minutes of the period.

Treatments

A description of the treatments follows:

1. *Programed Books:* The program was prepared on 5" x 8" cards, one card being used for each frame of the program, with the answer to the frame appearing on the reverse side of the card. The cards for each of the units were bound by loose-leaf rings. The resulting products were called "programed books." The subjects in this treatment studied a unit (one for each 75-minute class period) on a self-paced, but time-limited basis. Since the program required a written response for almost every frame, the students were provided with an answer sheet which corresponded, by frame numbers, with the unit being worked. Figure 1 shows a student at work. The photograph shows a programed book and the special answer sheet. Figure 2 shows the character of the "programed book" treatment as it was used in a class.

2. *Teaching Machines:* The entire program was also prepared on 8½" x 13" sheets of paper with the frames appearing in sequence, the answer to any given frame immediately

---

[1] First-period classes meet from 8:00 to 9:15 a.m.

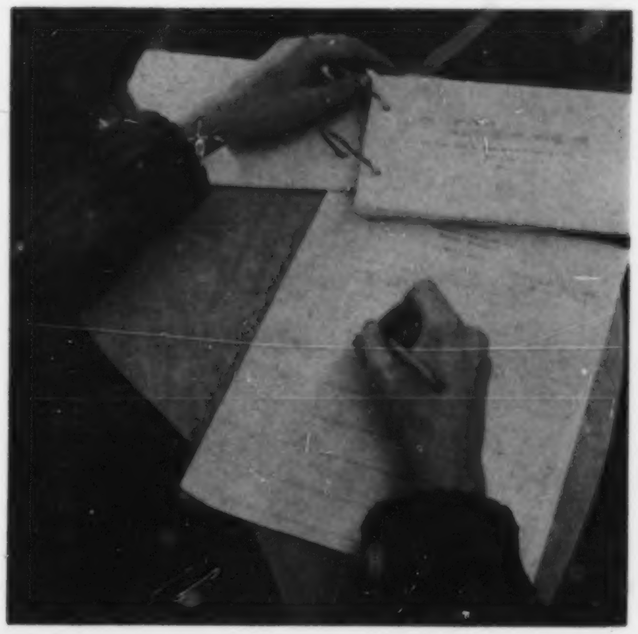


Figure 1. Student working on the programed book version of the algebra program.

Figure 2. Classroom showing use of programed books in the algebra program.

following the frame itself. All sheets required for any given unit were inserted in a simple teaching machine, an aluminum "envelope" which measured 2" x 10" x 17".[1] The characteristics of the "envelope" included a display window and a storage and receiving tray for the sheets of the program. The design of the machine required the students to move a sheet by hand from the top of the storage deck so as to display the items one at a time. The frame was moved into position, a response was made, and then the answer was moved to the display window. The completed sheet was reinserted under the delivery deck into a receiving tray. In this treatment (subsequently referred to as the "teaching machine" treatment) the students proceeded through the material on a self-pacing time-limited basis, completing one unit plus the unit-test each class period (75 minutes). The answer sheets used by the students in this treatment were identical to those used by the students in the "programed book" treatment.

Figure 3 shows an individual student working in the "teaching machine" treatment. It also shows a close-up view of the "envelope" and the format of the program for this treatment. Figure 4 shows the character of the "teaching machine" treatment as it was used in a class.

3. *Externally-paced Treatment:* For this treatment the entire program was copied on film strips with the frames appearing in sequence. A stimulus frame was followed by an answer or reinforcement frame.[1]

For the presentation, the frames were projected onto a 3′ x 4′ screen in the front of the room, and the students in this treatment (subsequently called the "film strip treatment") were required to respond to the stimulus frames on an answer sheet identical to that used in the "programed book" and the "teaching machine" treatments. The exposure times for the stimulus and answer frames were the mean times required for completion of the frames by the students who participated in the final testing of the program as described in Phase III, Development of the Programed Materials. Thus, the pacing rate in the film strip treatment was planned in such a manner that this group had approximately the same amount of time to complete each frame in each unit as that taken by the self-paced groups (programed "books" and "teaching machines"). Hence, the students in this treatment proceeded through the programed material—one unit each class period—as a group. This was the first test of the feasibility of the method and of the possibility of using programed materials with closed-circuit television. Figure 5 is a photograph which shows students in the "film strip treatment."

---

[1] The "envelope" was built by the Lock Haven Air Motive Corporation, Lock Haven, Pennsylvania, according to specifications designated by the researchers. A similar machine had previously been used by the Montgomery County Public Schools, Rockville, Maryland.

[1] Assistance was given by the Recordak Corporation, New York City, in producing the film strips.

4. *Conventional Class:* This treatment was not an integral part of the main comparison, but rather was arranged for the purpose of obtaining qualitative information about the program. An instructor[1] from the Mathematics Department agreed to teach the materials. In effect, she studied the programed units in advance in order to know their content and the manner in which the subject matter was developed. She was requested to teach the same content as that covered by the program using a lecture-discussion method familiar to her. It was suggested that she teach as she would normally, subject to the following restrictions: first, to teach one unit of material per class period, covering the same content that was taught under each of the other conditions; and, second, to provide her students only with those materials that were provided to all of the students in the other three treatments (viz., study guides).

Figure 6 shows the instructor teaching the class. A comparison of Figures 1, 3, 5, and 6 shows the similarity and differences among the four treatments. Figure 7 is a photograph of the conventionally-taught class taken during a regular class period. It is a picture of a teaching situation with which we are all familiar, and it is interesting to compare this photograph with those in Figures 2 and 4 from the point of view of the activity of the students involved.

Experimental Design and Procedures

An initial group of 113 students was randomly assigned to the four treatment groups described in the preceding section. Three of these groups were used for the main statistical comparisons (external-pacing vs. self-pacing), the fourth (teacher-taught class) being included for purposes of assessing program effectiveness. The subject matter taught in all of the treatments was the same. The method of presenting the materials was the experimental variable.

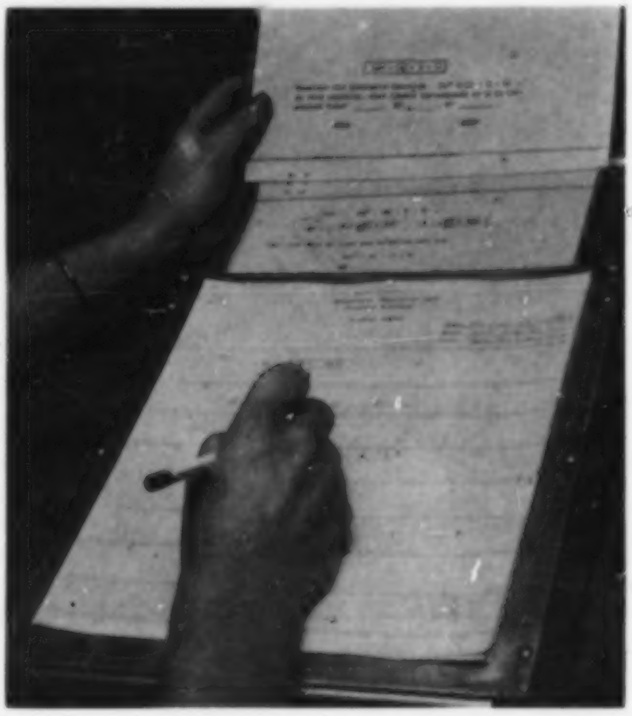


Figure 3. Student studying the teaching machine version of the program.

The experiment was run for a period of 15 class sessions of 75 minutes each, one unit of work being completed each session. A comprehensive examination was given during the sixteenth class period; moreover, a unit-test was administered at the conclusion of each class session. This latter procedure served to motivate the students, and to provide a test review of the unit just studied and an indication of each student's progress unit by unit. The comprehensive examination provided the same achievement measure for all four treatments.

Statistical comparisons were made on two criteria:

1) comprehensive examination scores

2) sum of the 15 unit-test scores

---

[1] Mrs. Hilda Duncan. She was selected by the Mathematics Department on the basis of her mathematical ability, her ability as a teacher, and her acceptance of the nature and content of the course that was programed.

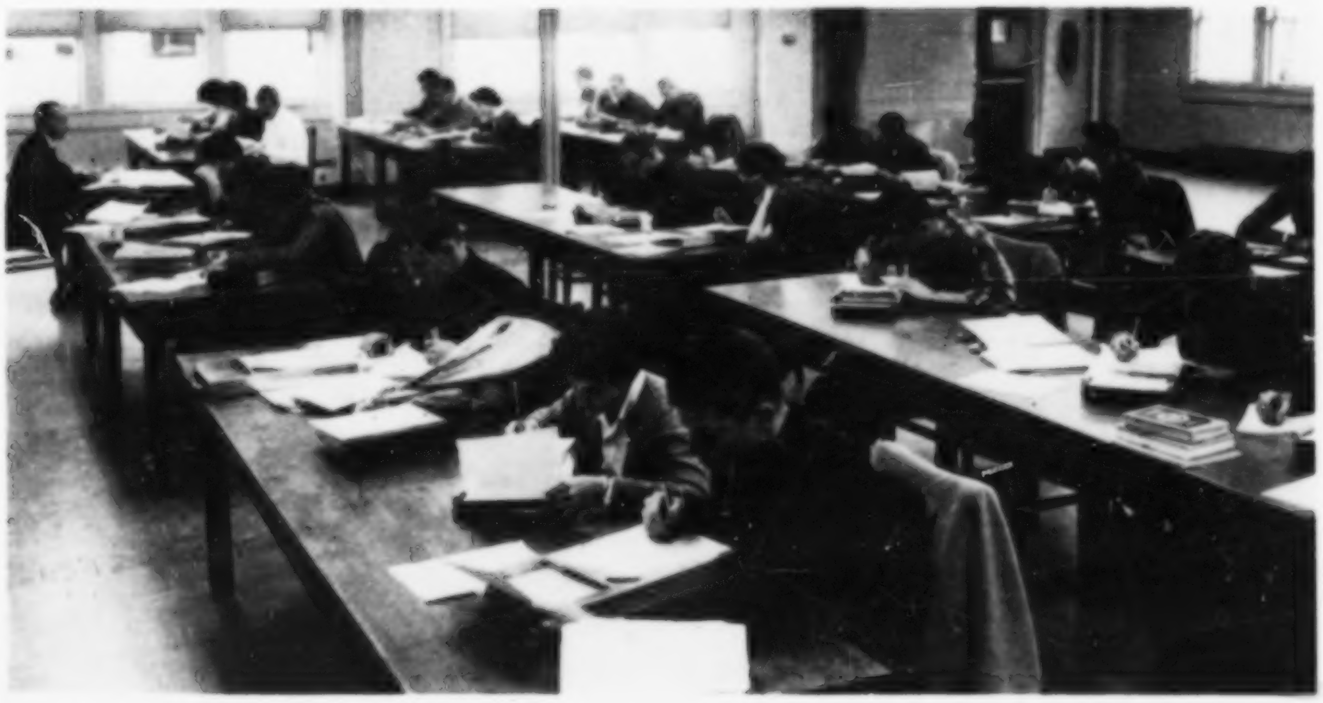


Figure 4. Classroom showing use of teaching machines.

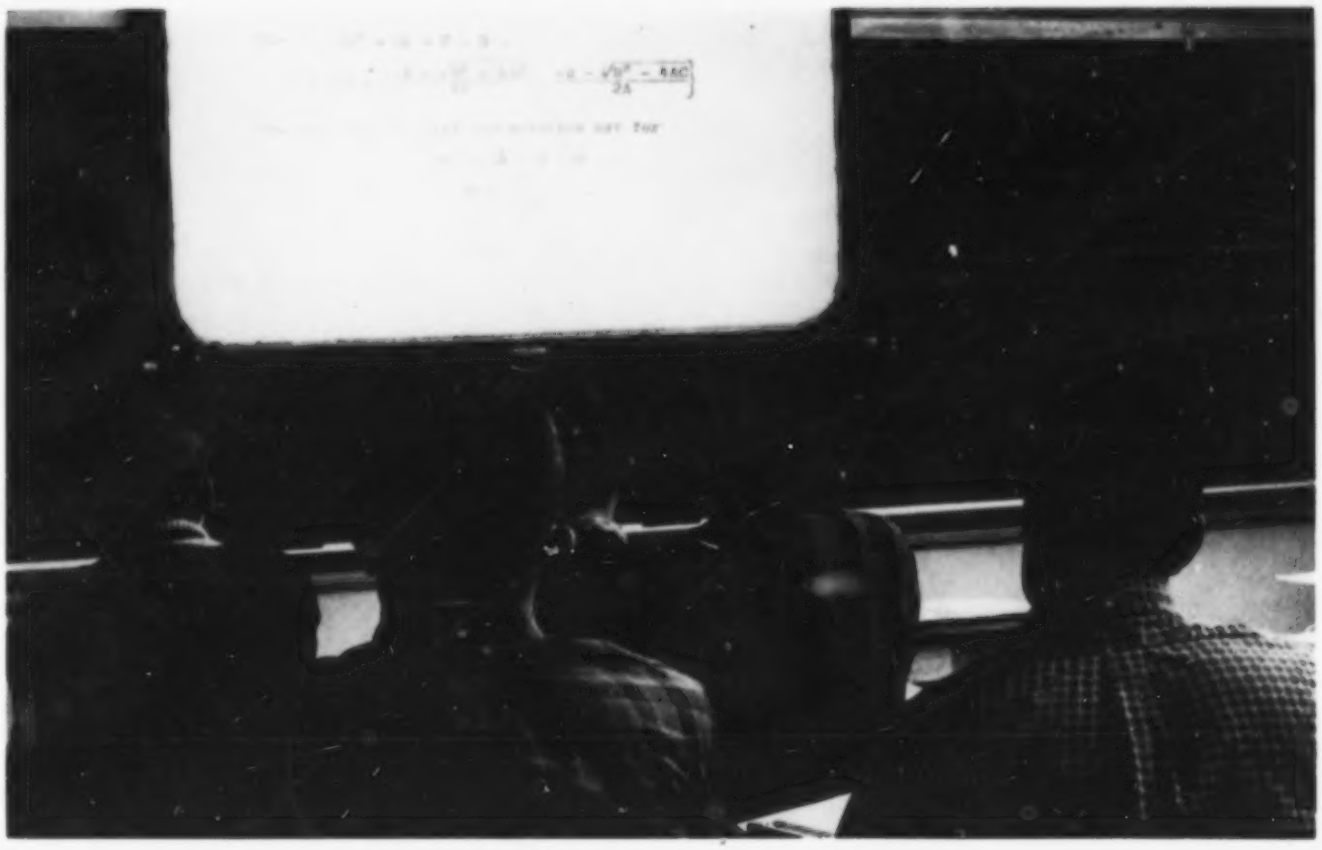


Figure 5. Students studying the externally-paced film strip version of the algebra program.

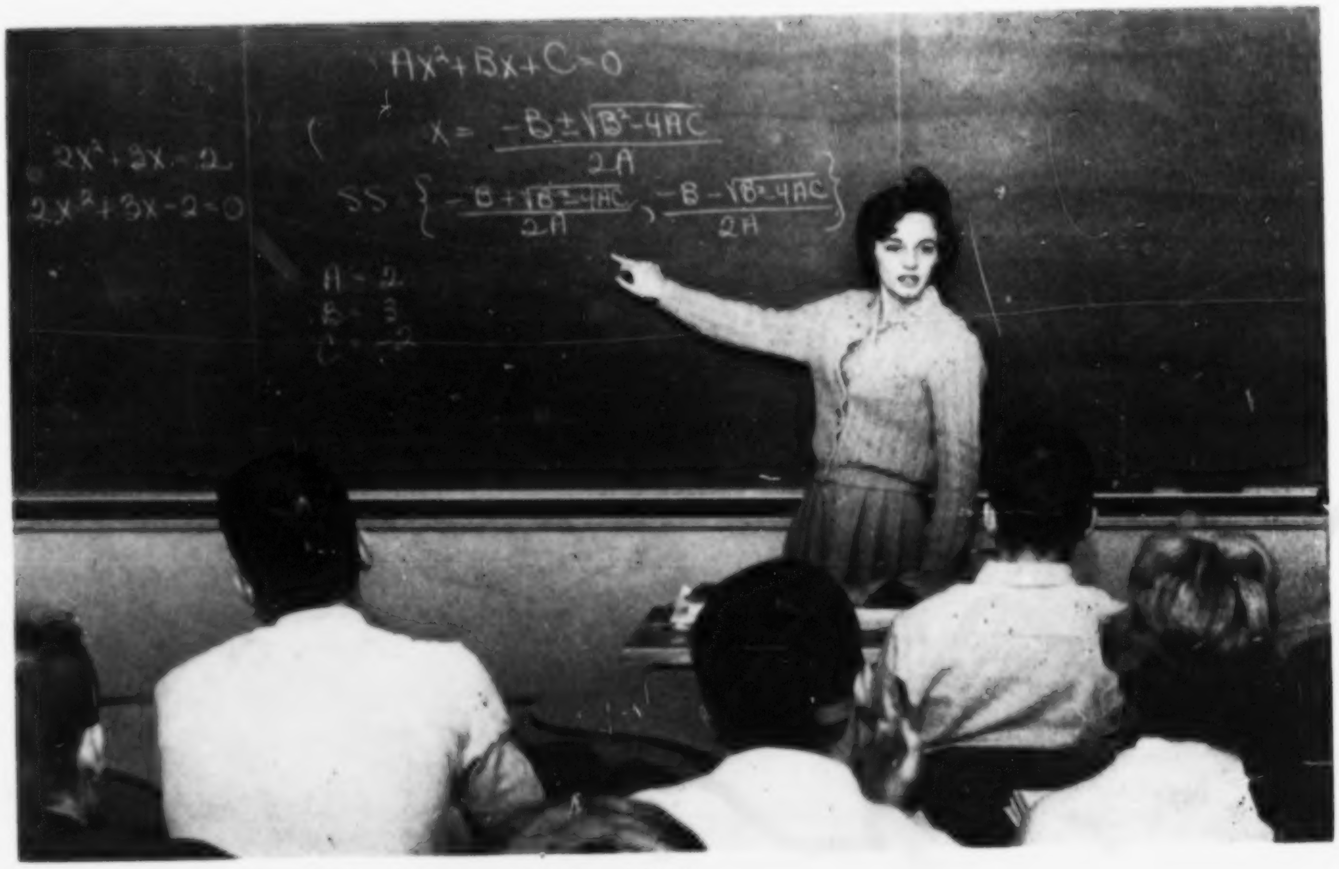

$Ax^2+Bx+C=0$
$x=\frac{-B\pm\sqrt{B^2-4AC}}{2A}$
SS $\left\{\frac{-B+\sqrt{B^2-4AC}}{2A}, \frac{-B-\sqrt{B^2-4AC}}{2A}\right\}$
$2x^2+3x=2$
$2x^2+3x-2=0$
A = 2
B = 3
C = -2

Figure 6. Class being taught by the lecture-discussion method.

Figure 7. Students during lecture-discussion presentation.

The effectiveness of the program was assessed by observing the percentage of test questions answered correctly and also by comparing the combined program treatments against the conventional treatment on each of the two criteria. Separate tests of significance were used to compare the means (t-test) and variances (F-test) with respect to each criterion.

For the main comparison (external-pacing vs. self-pacing), the analysis of variance method was applied to evaluate differences among program-treatment means on both criteria. In addition to these basic features of this experimental design described above, a number of administrative procedures were incorporated for "design" reasons:

1) At the first class meeting, all of the subjects involved in the experiment were instructed over television about the nature of programed materials, the experiment in which they would take part, the programed materials they would be studying, the basis for grading, and the specific requirements for success in the course. Upon completion of this orientation session, each student was informed of the treatment in which he would be working and was told to go immediately to the room where he, henceforth, would receive instruction. When the students were in their assigned rooms, they were given specific instructions about the particular treatment in which they would be involved. Also, they were given the opportunity to examine the materials they would be studying beginning with the second scheduled period.

2) At the conclusion of each unit, each student in each treatment was given the unit-test which corresponded to the unit of work for that day. In the "film strip" and the "conventionally-taught" treatments, the unit-tests were taken during the last 10 minutes of each period. In the cases of the "programed book" and "teaching machine" treatments, each student took the examination immediately upon completing the programed unit. (It should be recalled that the programed units were planned in such a way that they could be completed in an *average* time of 65 minutes.) To provide for the slow-working students in the "self-paced" ("programed books" and "teaching machines") treatments, students in these sections were given the privilege of coming to class and beginning work on the daily unit as early as 7:45 a.m. This procedure made it possible for all students to complete both the programed unit and the unit-test by 9:15 a.m.

3) The unit-tests were graded between successive class periods, and the scores were posted at the beginning of the next class. Therefore, each student was regularly informed of his achievements.

4) Study guides were distributed at the end of each period to the students in all sections, and these were retained by the students for study and future reference.

5) All students in all treatments received the same daily supplementary assignments. (All exercises were selected from those in Section VI of the study guides described previously.) The assignments were consistently graded in the period between classes, and the materials were returned to the students at the next class period. Thus, another kind of practice and knowledge of results was supplied to the students.

6) All students in all treatments took the comprehensive examination at the same time; namely, during the sixteenth regularly scheduled class period.

7) The Mathematics Department assigned a graduate assistant to each of the treatments. It was the responsibility of these graduate assistants to study the programed materials and be responsible for assigning final grades for the course. The graduate assistants also kept a "daily log" consisting of an absence record, a record of important events which occurred during the conduct of the various classes, and a record of the amount and extent of the outside assistance given to their students (this was negligible).

RESULTS

PART I: Effectiveness of the Program

The unit-tests and comprehensive examination were constructed to serve two measurement functions: first, to assess the extent to which the program actually taught what it was intended to teach; and, second, to provide data for the statistical comparisons among the different experimental groups.

Data on the first of these functions are presented in Table 4. This table gives, by unit, the percentage, respectively, of the unit-test items and the comprehensive examination items which were correctly answered by each of the four treatment groups. The over-all percentage of correct answers, on the criterion measures, of the three program treatment groups taken collectively is also given.

Two items of information about the effectiveness of the programed material may be gained from Table 4. Without making any direct comparisons, observation of the first three columns shows that at least one of the program treatments exceeded a performance level of 75 per cent on 13 of the 15 unit-tests; a comparable result was given by the comprehensive examination. The attained performance level of at least one of the treatments was greater than 80 per cent on four of the unit-tests.

When comparisons are made, only three of the entries in Column Five for the conventionally-taught class show an attained performance level in excess of 75 per cent, and none are above 80 per cent. The entries in Column Five are lower than all other corresponding entries with the exception of that for Unit VI. A comparison of the percentages given in Columns Four and Five is especially revealing. Figure 8 is a graphic representation of the percentages listed in Columns Four and Five which show the relationship between the corresponding entries.

In order to give some substance to the comparisons of interest, four tests of significance were conducted: one comparing the mean performances on the comprehensive examination of the conventional class with the program groups, taken collectively; and one comparing the variances of these groups on the same examination. Similar tests of significance were also conducted using scores based on the unit-tests. For the latter two comparisons, the various unit-tests were treated as one large examination, the score for any particular student being the sum of the scores he received on the unit-tests.[1]

TABLE 4

A TABLE OF THE PERFORMANCE LEVEL (IN PER CENT) OF THE SEVERAL TREATMENTS ON THE VARIOUS CRITERION MEASURES

Unit	Programed Presentations				Lecture Discussions
	Book	Machine	Film strip	Over-all	Conventional Class
I	.77	.78	.75	.77	.72
II	.78	.73	.72	.75	.66
III	.79	.78	‡	.79†	.69
IV	.81	.80	.78	.80	.73
V	.73	.73	.68	.72	.65
VI	.73	.75	.76	.74	.78
VII	.63	.63	.61	.63	.60
VIII	.78	.76	.73	.76	.69
IX	.78	.75	.72	.75	.69
X	.83	.81	.81	.82	.78
XI	.83	.81	.75	.80	.70
XII	.79	.73	.72	.75	.52
XIII	.71	.72	.74	.72	.69
XIV	.80	.83	.81	.82	.76
XV	.76	.77	.75	.76	.71
Final	.75	.77	.74	.76	.71

‡The third unit-test was not administered to the film strip group.

†This figure is based on the two treatments for which complete data were available.

Analyses Based on the Comprehensive Examination Scores

Table 5 provides statistics which describe the performance on the comprehensive examination of the program groups, taken collectively, and of the conventional class. Four statistics are given for each group: (1) the group mean, (2) the standard deviation, (3) the group variance, and (4) the number of subjects in the group.

---

[1] Since the students in the film strip treatment did not take the examination for Unit III, this test was omitted from consideration.

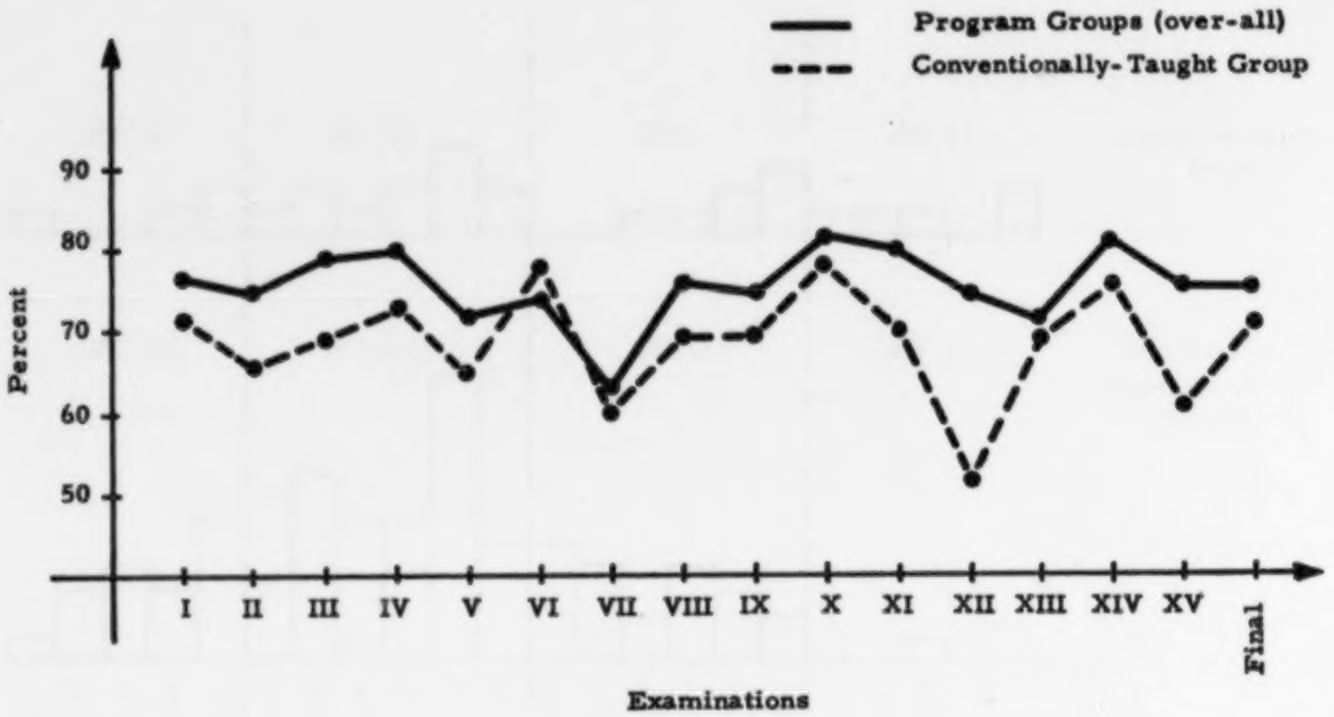


Figure 8.

Unit and Final Test Performances of the Conventionally-Taught Class as Compared with the Combined Program Groups

TABLE 5

THE MEANS, STANDARD DEVIATIONS, AND VARIANCES OF THE PERFORMANCES ON THE COMPREHENSIVE EXAMINATION OF THE COMBINED PROGRAM GROUP AND THE CONVENTIONAL CLASS

Combined Program Groups	Conventional Class
$\bar{X} = 45.35$	$\bar{X} = 42.57$
$s = 6.48$	$s = 8.26$
$s^2 = 41.93$	$s^2 = 68.15$
$n = 82$	$n = 28$

To test the hypothesis of equality of group means, the t-test was employed.[1]

A t-value of 1.62 was obtained. Since the critical value (at the .05 level of significance) for this test is 2.01, the hypothesis of equal group means could not be rejected.

To test the hypothesis of equality of group variances, the ratio of the sample variances which is known to have an "F" distribution was employed. An F-ratio of 1.625 was obtained. Since the critical F-value for this test was approximately 1.625, the hypothesis of equal treatment variances was rejected.

Analysis Based on the Summed Unit-Test Scores

Table 6 provides statistics which describe the performances of the combined program group and the conventionally-taught class on the unit-tests, taken collectively. As in Table 5, four statistics are given for each group: (1) the group mean, (2) the standard deviation, (3) the group variance, and (4) the number of students in the group.

TABLE 6

THE MEANS, STANDARD DEVIATIONS, AND VARIANCES OF THE PERFORMANCES ON THE SUMMED UNIT-TEST SCORES OF THE COMBINED PROGRAM GROUP AND THE CONVENTIONAL CLASS

Combined Program Groups	Conventional Class
$\bar{X} = 179.12$	$\bar{X} = 164.00$
$s = 20.13$	$s = 28.33$
$s^2 = 405.18$	$s^2 = 802.64$
$n = 82$	$n = 28$

---

[1] $$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

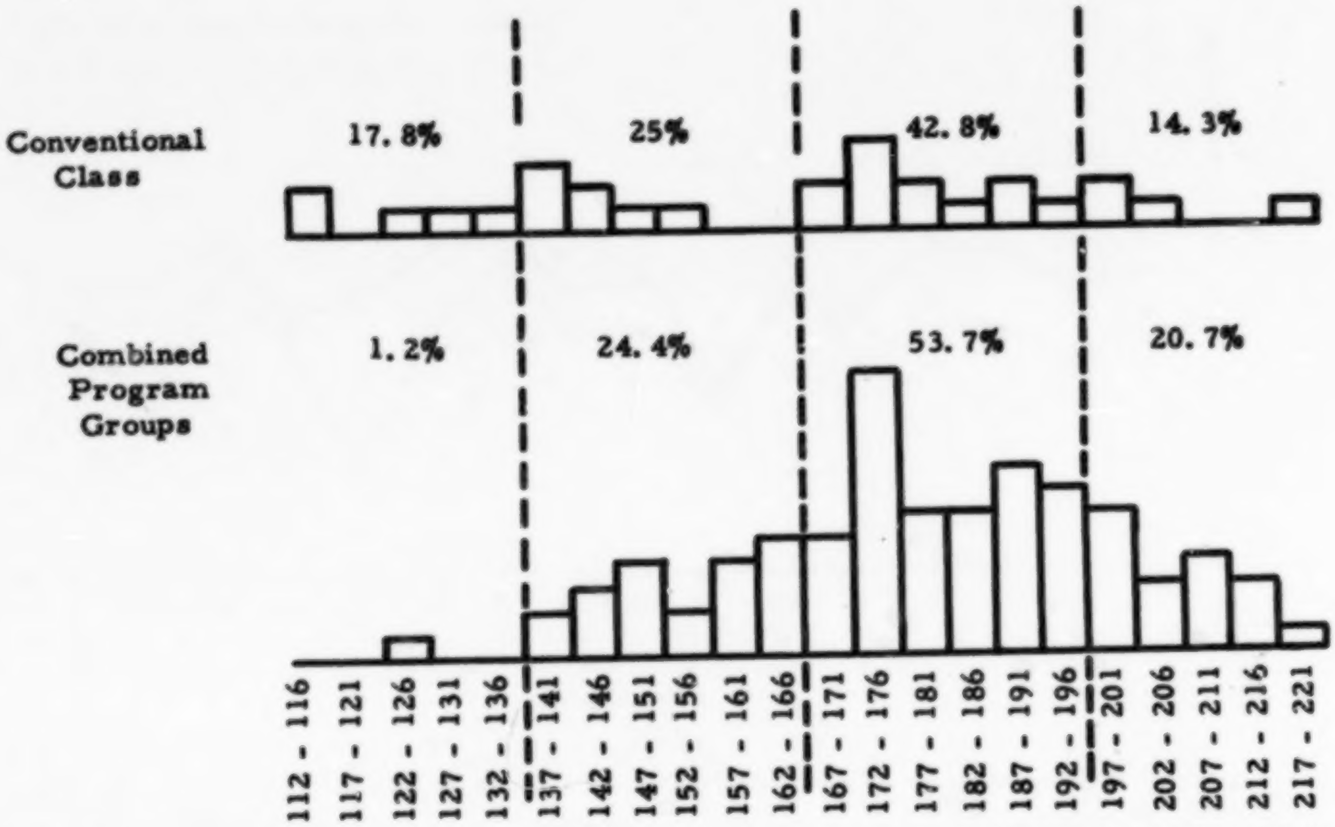


Figure 9. Histograms of the Frequency Distributions of the Summed Unit-Test Scores of the Conventional Class and the Combined Program Groups.

The t-test was again employed to test the hypothesis of equal means, and yielded a t-value of 2.61. Since the critical value (at the .05 level of significance) for this test was 2.01, the hypothesis of equal group means was rejected.

To test the hypothesis of equality of group variances, as before, the ratio of the sample variances was computed and yielded an F-value of 1.98. Since the critical F-value for this test was approximately 1.625, the hypothesis of equal group variances was rejected.

It is enlightening at this point to compare graphically the frequency distributions of the group performances. These frequency distributions are presented in histogram form in Figure 9. In order to provide additional information for the reader the total range of scores is divided approximately into quarters (as indicated by the broken lines) and the percentage of scores in each group which falls in each category is also given. An examination of Figure 9 reveals these facts:

1) The percentage of the scores of the program group which falls in the top quarter of the range is slightly larger than the percentage of the scores of the conventional class which falls into this category.
2) Only 1.2 per cent of the scores of the program group fall in the lower quarter of the range; in fact with one exception, all the scores in the lower quarter of the range belong to members of the conventionally-taught class.

These facts, in conjunction with the previously stated results, indicate that the significantly greater mean and lesser variance of the program-taught groups on the sum of the unit-tests is largely due to the almost total elimination of extremely poor performances by the program groups at no expense of top performances.

PART II: External-pacing vs. Self-pacing

A population of 113 students was originally available for the conduct of this research. These students were originally randomized among four treatment groups, one of which (the conventional class) was not employed for the main comparisons (external-pacing vs. self-pacing). Since 29 of the original 113 students were assigned to the conventionally-taught class, 84 students were initially involved in the main experiment. Of these students, one was lost through attrition over the treatment period, and one student was eliminated from the study at the discretion of the writer;[1] thus the remaining 82 students formed the experimental sample.

As was indicated earlier, the performances of the students (in the various treatments) on the criterion measures formed the basis of the analysis. Two analysis of variance tests were conducted; one used the scores on the comprehensive examination; the other used the sums of individual scores on the various unit-tests.

Analysis Based on the Scores on the Comprehensive Examination

Table 7 provides statistics which describe the performances of the three experimental treatment groups on the comprehensive examination. Four statistics are given for each group: (1) the group mean, (2) the standard deviation, (3) the group variance, and (4) the number of students in the group.

TABLE 7

THE MEANS, STANDARD DEVIATIONS, AND VARIANCES OF THE THREE PROGRAMED TREATMENTS ON THE COMPREHENSIVE EXAMINATION

Programed Book	Teaching Machine	Film Strip
$\bar{X}$ = 45.3	$\bar{X}$ = 46.3	$\bar{X}$ = 44.4
s = 10.78	s = 11.02	s = 11.15
s^2 = 116.25	s^2 = 121.40	s^2 = 124.47
n = 28	n = 28	n = 26

Before the analysis of variance was evaluated, Bartlett's test was applied to determine whether the requirement of homogeneous variances was satisfied. The obtained F-ratio of .012 (2,79 degrees of freedom) justified the decision that the true variances were homogeneous.

Since the variances were homogeneous in this instance, the analysis of variance test was applied. Essentially this test consists of assuming that the true treatment means are equal and, on the basis of the sample means which are obtained, making a judgment as to whether this hypothesis is tenable. The complete summary of the analysis of variance data for the scores on the final examination is given in Table 8.

TABLE 8

ANALYSIS OF VARIANCE OF THE COMPREHENSIVE EXAMINATION SCORES FOR THE THREE PROGRAMED TREATMENTS

Component	S.S.	df	V	F	p
Treatment	49.04	2	24.52	.58	>.05
Error	3361.71	79	42.55		
Total	3410.75	81			

On the basis of this analysis the differences in performances among the three programed treatments on the final examination are judged to be attributable to chance.

Analyses Based on the Summed Unit-Test Scores

Table 9 provides statistics which describe the performances of the three program treatment groups on the summed unit-test scores. As in Table 7, four statistics are given for each group: (1) the group mean, (2) the standard deviation, (3) the group variance, and (4) the number of students in the group.

TABLE 9

THE MEANS, STANDARD DEVIATIONS, AND VARIANCES OF THREE PROGRAM TREATMENT GROUPS ON THE UNIT-TESTS, TAKEN COLLECTIVELY

Programed Book	Teaching Machine	Film Strip
$\bar{X}$ = 182.00	$\bar{X}$ = 180.04	$\bar{X}$ = 175.04
s = 23.68	s = 27.99	s = 24.47
s^2 = 560.89	s^2 = 783.58	s^2 = 598.79
n = 28	n = 28	n = 26

---

[1] This decision was made on the basis of the fact that the over-all performance level of this student on the 15 unit examinations was approximately four standard deviations below the mean performance level of his group. Moreover, his score on the final examination was approximately 2.7 standard deviations below the treatment mean score.

Bartlett's test was again applied to compare sample variances for the summed unit scores. The requirement of homogeneity of variances was met in this instance also.

The complete summary of the analysis of variance data using the summed unit-test scores is given in Table 10. The hypothesis being tested is, again, that of equal treatment means.

TABLE 10

ANALYSIS OF VARIANCE DATA FOR THE SUMMED UNIT-TEST SCORES FOR THE THREE PROGRAMED TREATMENTS

Component	S.S.	df	V	F	p
Treatment	688.86	2	344.43	.84	$>.05$
Error	32509.93	79	411.52		
Total	33198.79	81			

As in the previous analysis, the differences in performances among the three program treatment groups on the unit-tests, taken collectively, were judged to be due to chance.

DISCUSSION OF RESULTS

Effectiveness of the Program

Taking into account the average performance (over 75 per cent of the test items answered correctly) and the level of understanding demanded by those questions, the program was judged to have attained a very satisfactory level of teaching effectiveness.

Comparisons of the combined program treatments and the conventional treatment were conducted on means and variances to gain further evidence bearing on the effectiveness of the program. For the comprehensive examination, the statistical tests revealed no significant difference in mean performance, but the variances were found to be significantly different, the smaller variance belonging to the program groups.

For the combined unit-tests, the statistical analyses revealed significant differences in both means and variances, the greater mean and lesser variance belonging to the program groups. A histogram comparing the frequency distributions of the program group and the conventional class indicated that this combination of difference worked in favor of the program group, the decrease in variance being due largely to fewer extremely poor performances in the program group (at no expense of top performances) than were manifested by the conventionally-taught class.

Thus, the statistical tests comparing performances of the combined program groups and the conventional class produced some evidence to justify the judgment that the program was an effective teaching instrument.

In summary, the results described above support the conclusion that the program which was prepared as an integral part of this study is an effective instrument for teaching what it was intended to teach—as measured by the several criterion examinations.

External-pacing vs. Self-pacing

Of the three program treatments whose effects were assessed in the main comparison, the "film strip" presentation was the most interesting, because it was anticipated that this course would be adapted for use over television. The group receiving this treatment was paced at a rate where the time allowed for completion of any given item was the average time required for the completion of that item by a group of self-paced students who were involved in the pre-experimental testing of the program.

Students in the "programed book" and "teaching machine" treatments proceeded through the programed material on a self-paced but time-limited basis.

All 15 programed units were completed on the basis of one unit per class period; and, in all treatments, unit-tests were given immediately after the students finished the units in question. Thus, 15 class periods were needed to complete the entire sequence of units, the sixteenth class period being employed for the purpose of administering the comprehensive examination. Two tests were conducted, one using the comprehensive examination scores and the other using the summed unit-test scores. The results on both of these criteria supported the hypothesis of equality of group means; that is, no significant differences among the treatments were found in either case.

CONCLUSIONS

The results of the two analysis of variance tests which were conducted indicate that the effects of the three treatments under consideration were essentially equivalent. In view of this fact, it seems reasonable to conclude that the use of programed materials for purposes of *group* instruction under externally-paced conditions is feasible.

A corollary conclusion is in order: one of the basic tenets of programed instruction suggests that the most advantageous condition for learning is one which permits each individual to control his learning pace. Since the results of the research reported here do not substantiate this assumption, it seems reasonable to conclude that the assumption may need to be qualified. The results of these comparisons increase confidence in the possibilities of reciprocally adapting programed courses and closed-circuit television for instructional purposes.

EXPERIMENT II: An Experimental Comparison of Differential Rates of Pacing Programed Mathematics

by J. J. Lottes, Gilbert Palmer, and Herbert Oakes

PURPOSE

During the Spring Term of 1962, a study was conducted to explore the effects on the learning of students of different rates of pacing the presentation of programed materials.

The purpose was two-fold:

1) to provide information relative to the optimal rate of presentation of programed materials by means of the mass media

2) to gain evidence bearing on the validity of the assumption that freedom for each student to work at his own rate is an essential concomitant of effectively utilized programed materials

To satisfy the above stated purposes, an experiment was conducted at State University of New York, College of Education at Geneseo,[1] to compare the effects of varying the rates of presentation of programed modern algebra.

Statement of the Problem

It was intended to compare experimentally the efficacies of varying rates of pacing the film strip presentation of a part of the programed course in Contemporary Algebra.

Comparisons were made with respect to two criteria:

1) achievement, as measured by unit achievement tests
2) attitude toward the method of instruction, as measured by a unidimensional scale

A second problem was to compare experimentally the effectiveness of programed modern algebra presented under externally-paced conditions against the identical programed material presented under self-paced conditions.

METHOD

Subjects

The subjects were 180 Geneseo College students enrolled in General Mathematics I, a course required of all students except for a small proportion exempted as a result of screening by a proficiency examination.

Materials

The programed materials in Contemporary Algebra described in an earlier section of this report, were used as the academic content for the experiment. The first six units of the programed course were used; these units were:

Unit I—Numbers
Unit II—Sets
Unit III—Sets and Variables

---

[1] Principal investigators: J. J. Lottes, G. Palmer, and Herbert Oakes of State University of New York, College of Education at Geneseo. The research was sponsored by the Course Development Project at Penn State.

Unit IV—The Language of Algebra
Unit V—Operations; Part I
Unit VI—Operations; Part II

The Study Guides, which provided summaries of each unit and appropriate practice exercises (for assignment), were also used.

A unit-test (14-19 items), also previously described, was administered at the completion of each programed unit. Comparisons of achievement used the total of the six unit-tests as the criterion score; the Kuder-Richardson Formula 20 reliability computed for this criterion was .70.

Attitudes were measured on a unidimensional scale, *Attitude Toward Method of Instruction*, developed by Frank Hartman at The Pennsylvania State University.[1] The obtained reproducibility was .93, which may be interpreted as a guarantee of high reliability.

Experimental Treatments

The treatments consisted of four different rates of presentation of the programed algebra presented by film strip projection. The times for each frame were 80, 90, 100, and 110 per cent of the base times; the base times were those determined in a pilot study at The Pennsylvania State University and were the mean time for each frame computed on a group of 14 subjects (individually-timed), each of whom proceeded at his own rate under self-pacing conditions.

In addition, two groups were entirely self-paced; the programed material was presented in book form, and each student proceeded through the material at his own rate.

Experimental Design

A population of 180 subjects was randomized over the six treatment groups; four of these groups (externally-paced) were used for the main purpose of comparing varying rates of film strip presentation, the other two groups (self-paced) were used only for the secondary purpose of comparing self-paced presentation with externally-paced presentations.

The six programed units of Contemporary Algebra were administered to each treatment group, with a maximum of two 50-minute class sessions being allotted for each unit. A uniform orientation was provided for all groups.

Upon the completion of each unit, a unit-test was administered (14-19 items). The achievement criterion was the sum of the six unit-test scores. After the unit-test had been taken, study guides were distributed and assignments were made.

After the completion of the last unit-test, all subjects responded to a scale, **Attitude Toward Method of Instruction,** and also to a questionnaire which provided an opportunity for students to write their comments and criticisms.

The main comparisons were made to evaluate the effects of varying the rate of externally-paced presentation of the programed materials. Analysis of covariance was applied to evaluate the treatment effects on the achievement criterion. The Geneseo College Mathematics Proficiency Examination was used as the adjusting variable. The comparison of means on attitude toward the method of instruction was conducted by applying the analysis of variance.

For the second problem, a t-test was used to assess differences between the externally-paced and self-paced treatment groups with respect to achievement.

RESULTS

Comparison of Varying Rates of External-pacing of Presentations on Achievement Measure

An analysis of covariance was conducted to assess differences among sample means. Results of the routine tests for significance of regression, homogeneity of regressions, and homogeneity of variances indicated that the requirements for application of the covariance model were satisfied. Consequently, the covariance analysis was carried out, for which the summary data appear in Table 11, below.

The F-ratio of .58 (3,111 degrees of freedom) indicated that the differences among means were not significant at the .05 level. Therefore, in this experiment, the variations of

---

[1] Hartman, Frank R., *The Assessment of Respondent Error in Attitude Scaling.* Unpublished doctoral dissertation, The Pennsylvania State University, 1957.

TABLE 11

ANALYSIS OF COVARIANCE FOR ACHIEVEMENT OF GROUPS RECEIVING VARIED PACING OF PROGRAMED MATERIALS

Component	S.S.	df	V	F	p
Treatment	73.23	3	24.41	.58	p>.05
Error	4682.36	111	42.18		
Total	4755.59	114			

pacing rates used produced no significantly differential effects on the achievement measures.

The summary of observed and adjusted means is shown in Table 12, below. It should be noted that the achievement criterion had a possible score of 97.

TABLE 12

OBSERVED AND ADJUSTED MEANS FOR ACHIEVEMENT

		Achievement	
Rate of Presentation (Per cent of Base Times)	Pre-test Mean	Observed Mean	Adjusted Mean
80%	20.4	73.5	73.2
90%	20.1	73.7	73.6
100%	20.3	75.5	75.3
110%	19.1	73.9	74.4

Attitudes Toward Method of Instruction

Attitudes toward method of instruction were compared through application of the analysis of variance model. The F-ratio of 1.40 (3,108 degrees of freedom) indicated that the differences among population means were not significant at the .05 level. Therefore, the variations of pace produced no significant differential effects with respect to attitude toward the method of instruction. The summary data are displayed in Table 13.

Obtained group means are shown in Table 14 (high score is associated with favorable attitude).

TABLE 13

ANALYSIS OF VARIANCE FOR ATTITUDE TOWARD METHOD OF INSTRUCTION

Component	S.S.	df	V	F	p
Treatment	320.31	3	106.77	1.40	p>.05
Error	8245.55	108	76.35		
Total	8565.86	111			

TABLE 14

OBSERVED MEAN SCORES ON ATTITUDE TOWARD INSTRUCTION

	Treatment (% of Base Times)			
	80%	90%	100%	110%
Mean	32.8	30.8	31.1	28.2

External-pace Versus Self-pace

Although the experimental design was intended to include a comparison of the group (externally-paced) presentation against the individual (self-paced) presentation, it was recognized that a completely new experiment would have to be conducted to insure a valid probability statement. However, it was decided that some substantial information could be gained from such a comparison, particularly since it replicated Experiment I.

A t-test was conducted to compare the externally-paced and self-paced treatment means on the summed unit-test criterion. The obtained value of 1.66 (1,178 degrees of freedom) indicated that the difference between means was not significant at the .05 level. Therefore, it must be concluded that the externally-paced (film strip) and self-paced (programed books) methods of presenting programed algebra were equally effective.

The means and standard deviations are shown in Table 15.

TABLE 15

OBSERVED MEANS AND STANDARD DEVIATIONS FOR ACHIEVEMENT

	Externally-paced (% of Base Rate)				Self-paced	
	80%	90%	100%	110%	group 1	group 2
Mean	73.5	73.7	75.5	73.9	76.4	75.9
Standard Deviation	8.18	7.23	6.52	8.08	7.18	7.73

DISCUSSION OF RESULTS

Varying Rates of Externally-paced Presentations

Six units of the programed course in Contemporary Algebra were presented to four treatment groups by means of film strip projection, the independent variable being the rate at which the frames were presented. The times allowed for presentation of each frame were 80, 90, 100, and 110 per cent, respectively, of the base times; the base times were determined by using the mean frame-times obtained from individually-timed students.

Analysis of the performances of the groups on the summed scores of six unit-tests showed no significant differences, a finding which was rather surprising to the investigators since pilot-study experience had indicated that the 80 per cent time forced most students to work at the maximum limits of their reading and response speeds. Two reasons are suggested that may account for the lack of differences:

1) Repetition built into the program provided an opportunity for the faster-paced groups to learn the material adequately, even though concepts may not have been understood at first.

2) Individuals have rather wide *tolerance limits* for variations of the pacing of their learning.

With respect to the second of the above speculations, the classroom observations of the investigators are instructive. It was noted that the groups which were relatively slowly-paced tended to utilize their "spare" time, not to concentrate on the mathematical activity, but to relax until the next frame appeared on the screen. On the other hand, the fastest-paced students were highly attentive during practically 100 per cent of the presentation period. The investigators were impressed with the degree of persistent and intensive effort shown by the latter students.

Both of the factors suggested above lead to the conclusion that for the mathematics program that was used in the experiments, the rate of externally-paced presentation is *not* a critical factor within the limits tested relative to the achievements of students.

It was interesting also to observe that there was no significant difference among means on the criterion of attitudes toward the pacing rate variations with the programed method of instruction. Informal questionnaire responses also indicated that students reacted favorably to the method of instruction regardless of the rate of presentation of the content.

It is believed that the rate corresponding to 80 per cent of base time approaches a practical limit of speed, but further experimentation is needed to learn what are the ultimate limits.

External-pacing Versus Self-pacing

Although use of the external pacing groups for a second comparison was of some consequence relative to the validity of the probability statement, it was judged that the error would not be so great as to vitiate the information gained. Therefore, a t-test was applied to test the hypothesis of equal treatment means on the achievement criterion. It was concluded that the externally-paced and self-paced techniques of presenting programed materials were equally effective.

It should be noted that this conclusion supports Experiment I of this report.

CONCLUSIONS

The results of these experimental comparisons make tenable only one broad conclusion: for the content and programs used, and within the range of rates tested, the rate of presentation is *not* a critical factor. The capacity of the individual to perform at different rates of presentation, externally-paced or self-paced, is a factor that can be varied over considerable ranges without affecting learning as measured.

Although the self-pacing assumption about the presentation of programed materials has had almost universal support, the results of this study further support the conclusion that pacing within limits is not critical to program effectiveness. It would appear that college students can tolerate rather wide variations in performance rates. Furthermore, attention should be given to two concepts stated in

Part I of this report; namely, that the prevailing self-pacing rate of a student may not be the best rate, and that a progressive increase in pacing rate may improve the learning efficiency of a student working with programed materials. Therefore, presentations of programed materials by means of mass media may be reasonably developed without violation of a valid theoretical property.

EXPERIMENT III: Adaptation and Evaluation of a Programed Mathematics Course for Televised Presentation

by L. P. Greenhill, J. J. Lottes and A. Pagano

PURPOSE

The main purpose of this research and development project was to answer the question of whether programed learning can be mediated by television, and if so, with what degree of effectiveness. Orientation to this problem was given in Part I of this report. Experiments I and II, presented in the preceding sections of this report, were planned and conducted, in part, to investigate basic assumptions about methods of presenting programed instruction and especially pacing rates, which require regulation in order to present programed materials to groups by television. It was for the purpose of adapting programed materials for presentation over a television system and for evaluating the resulting effects on learning that Experiment III was planned and conducted.

STATEMENT OF THE PROBLEM

The problem was two-fold:

1) to adapt the programed course "Contemporary Algebra" for presentation over closed-circuit television

2) to evaluate experimentally the effects on learning of the televised presentation of the programed materials

The experimental assessments were accomplished by comparing the televised presentations with individual teaching machine presentations of the same programed materials, on the criterion of achievement as measured by unit-tests, midterm, and final examinations.

The problem of adapting the programed course to television will be discussed first; this discussion will be followed by a description of the experimental comparison of the two treatments.

ADAPTATION OF THE PROGRAMED COURSE FOR TELEVISION PRESENTATION

The following is a brief review of the characteristics of the program which was adapted to television:

1) This study used 28 units of the programed course "Contemporary Algebra." Each unit had been designed to be completed in approximately 65 minutes, thus allowing time within a 75-minute class period for administration of a unit-test.

2) The program was of the Skinnerian "linear" type which required constructed responses. Each unit was designed to include a review of the previous unit, a presentation of new programed material, and a review of the unit, respectively.

Detailed descriptions of the program and its development are given in an earlier section of this report.

Phase I of Adaptation to Television

The initial adaptation of the course for televised presentation (Unit I only) was similar, with some modification, to the film strip version used in Experiment I. This version consisted of typed material on 5" x 8" cards. These cards were referred to as "visuals." Because of the restricted area of the TV screen and the distance from which the students viewed the screen, the amount of print was limited to 11 lines of 26 characters each. (Students viewed the 23" television screen from a maximum distance of 12 screen widths). The print characters were typed in an area which had the best aspect ratio for televising them for readability in the classroom. Frames that contained too much material for this area were divided into two or more subframes and typed on separate cards; however, care was taken to insure that the content was not disordered.

One major difference between the first television adaptation and the film strip version of the program was the use of the auditory channel in the television version. The use of sound in this version served two functions:

1) to supplement the information presented in a frame on the screen, and
2) to read some or all of the material that appeared on the screen in a given frame. This second use of the audio channel enabled the students to hear as well see the symbols and words used. Appropriate emphasis could also be provided.

Major decisions that had to be made related to the question of how much commentary should be incorporated and how this was to be done. Versions in which the auditory channel was used to varying degrees were tried out experimentally.

An illustration of one of these trial forms follows; in this example the audio channel was used much more extensively than in most of the final television version.

Example #1

A. *Original version of frame:* 3

There is quite a difference between writing
1, 2, 3, 4
and
1, 2, 3
On the other hand is there a difference between writing 1, 2, 3, 4, . . .
and
1, 2, 3, . . . ? .... (yes or no)

STOP STOP

- - - - - - - - - - - - - - - - - -

no 3

B. *Television version of frame:*

	VIDEO	*AUDIO*
Visual #1	1, 2, 3, 4 is different from 1, 2, 3.	Writing 1, 2, 3, 4 is different from writing 1, 2, 3.
Visual #2	Is 1, 2, 3, 4, . . . different from writing 1, 2, 3, . . . ?	On the other hand, is 1, 2, 3, 4, etc. different from 1, 2, 3, etc., when the string of dots is used. Answer "yes" or "no."
Visual #3	no.	The answer is no.

Another method of presentation of the visual programed material involved the development of a display board with pull tabs, which enabled successive steps in the program to be revealed. In addition to the trial-methods described above, another technique of presenting the program with television was tried. This technique consisted of:

1) modifying the programed material for intermittent sequences of live "programed lecture" presentation
2) giving a demonstration "programed lecture" under simulated classroom conditions and obtaining a sound recording of this lecture
3) preparing visuals to be presented in synchronization with the sound tape

Phase II of Adaptation to Television

The experience gained during the first exploratory phase led to the decision to use the printed programed material as the main substance of the televised version, augmented by the periodic appearance of the instructor on television for introducing new topics and for review. In addition, the instructor's voice was used at appropriate points for emphasis or clarification of the programed presentation or for reinforcement.[1]

A brief description of actual procedures used in presenting a unit by television follows:

1) The typical procedure was for the instructor to appear on television to review the previous unit and introduce a new unit. He also would present a summary at the end of the unit or sequence, and explain parts of the unit of study. (We will refer to this use of the instructor as "instructor presentation.") Instructor presentations were made about three or four times per unit, and constituted about 15 to 20 per cent of the total presentation time. The instructor viewed a monitor screen while seated at a desk and used previously prepared visual materials to illustrate his presentations. These materials were prepared on blue paper placed in acetate folders. The instructor could write on the acetate folders with a grease pencil. These materials were televised by means of an overhead mirror arrangement. (See Figure 11.) The folders were arranged in sequential order to correspond with the instructor's presentation which was written in script form.

2) The remainder of the unit consisted of sequences of programed materials, for which the television system served as the teaching machine. The receivers in classrooms corresponded to the display areas in teaching machines. That is, programed material appeared on the screen for students to read. Except for occasional information on an orientation frame, each frame required the student to respond overtly on an individual answer sheet. At a pre-determined time, the answer frame was displayed and a chime was used to signal the appearance of a new frame on the screen. For the most part, mean item completion times obtained from individuals progressing on a self-paced schedule determined the rate at which frames were displayed. Audio comments were used occasionally to provide the answer to a question, or to supplement the printed question frame and/or answer, and to emphasize a point. These audio comments were written in script form and were displayed to the instructor by means of a television prompter. (See Figure 11.) Frame numbers did not appear on the television screen. However, from time to time the instructor would say, "you should write your response in space number .... on your answer sheet." Two examples illustrating the technique of adapting the original printed program for television presentation follow.

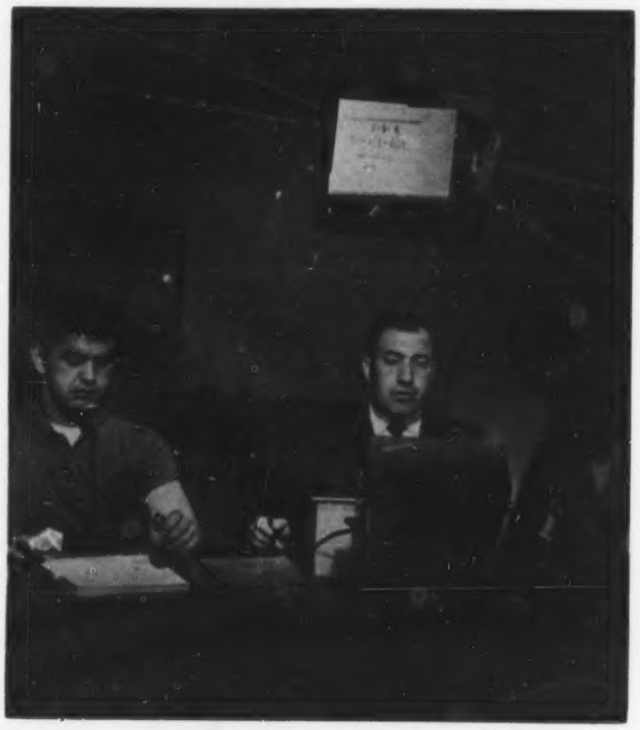


Figure 11. Studio arrangement for televising the instructor presentation sequences.

A. *Original version of frame:* 36

Every real number n, $n \neq O$, has a multiplicative inverse which we denote by $\frac{1}{n}$. That is, the multiplicative inverse of 7 is $\frac{1}{7}$. The multiplicative inverse of 4 is .... since 4· .... = 1.

---

[1] The instructor was A. Pagano.

STOP STOP

- - - - - - - - - - - - - - -

CORRECT ANSWER: 36.

$\frac{1}{4}$ in each case

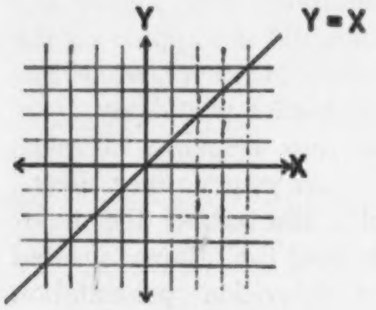
B. *Televised version of frame:*

VIDEO	*AUDIO*
Visual #1 Every real number n, n ≠ O, has a multiplicative inverse which we denote by $\frac{1}{n}$.	no audio
Visual #2 That is, the multiplicative inverse of 7 is $\frac{1}{7}$. The multiplicative inverse of 4 is .... since 4· .... = 1.	no audio

CORRECT ANSWER:
"$\frac{1}{4}$ in each case."

Figure 12.

A. *Original version of frame:* 71.



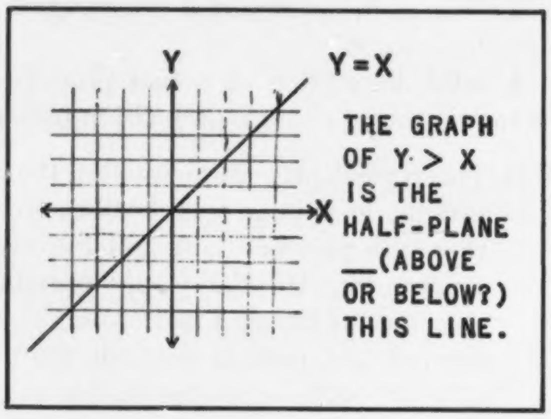
This line represents the graph of $y = x$. We have learned that the graph of $y > x$ is the half-plane ...... (above or below?) this line.

STOP STOP

- - - - - - - - - - - - - - -

CORRECT ANSWER: 71.
above

B. *Televised version of frame:*


Y
Y = X
X
THE GRAPH OF Y > X IS THE HALF-PLANE __(ABOVE OR BELOW?) THIS LINE.

Audio
"This line represents the graphing of $y = x$."

CORRECT ANSWER:
"above"

Figure 13.

Figure 14 is a photograph illustrating how a programed sequence appeared to the student; this photograph was taken during a regular class session.

Steps Followed in Preparing Programed Materials For Television

In order to modify and adapt each unit for televised presentation the following procedure was used:

1) A mathematics programer critically analyzed each of the programed units and, using the information gained during the program development, determined those parts which could be handled by the programed sequences alone and those parts which might be improved by using the instructor on television.
2) The frames to be included in the programed sequences were then prepared for televised presentation by typing the visually presented information on 5" x 8" cards. An IBM electric typewriter was used and the typing was confined to an area of $2\frac{3}{16}$" x $2\frac{13}{16}$". Frames which required more than the space available were subdivided into

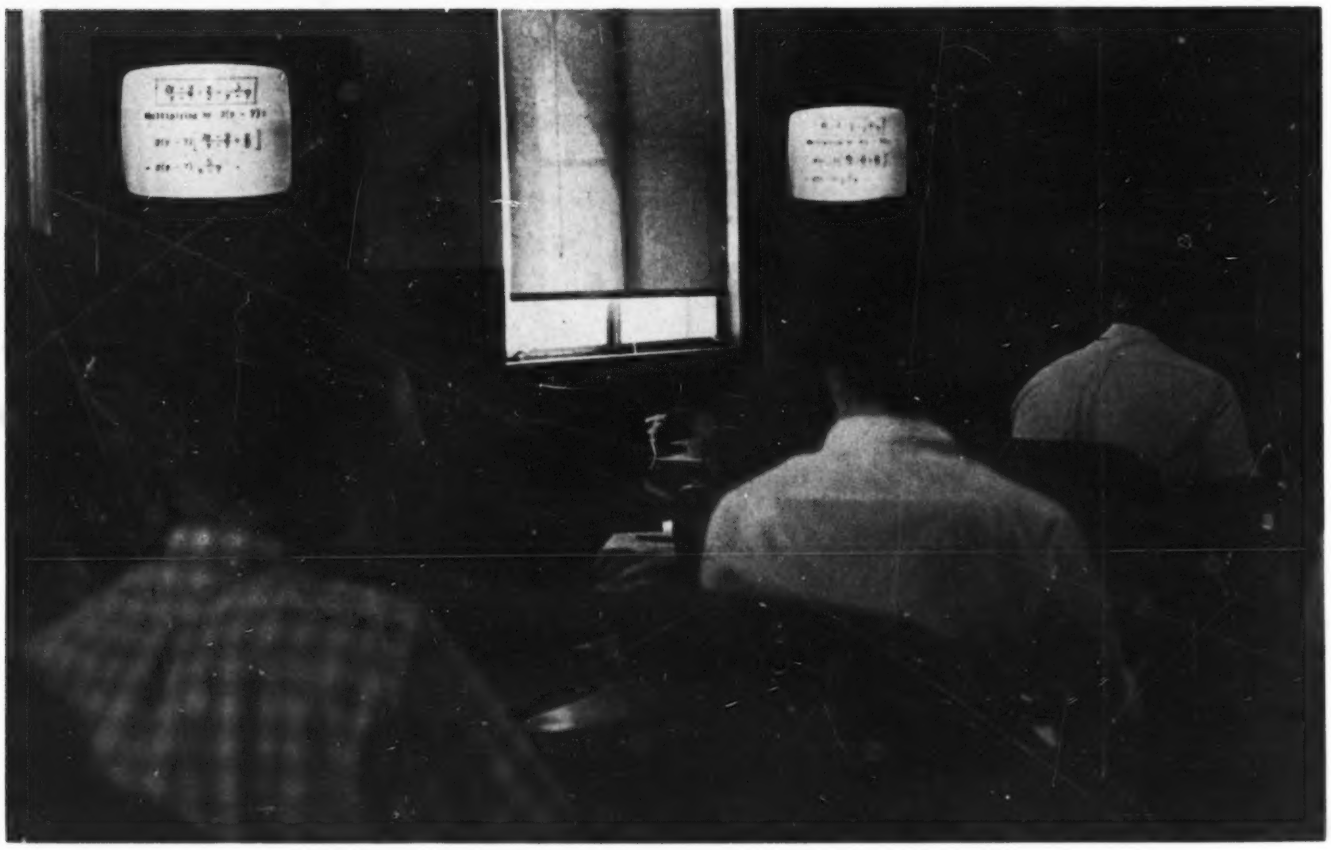


Figure 14. Television classroom showing students studying programed sequence in Contemporary Algebra course.

two or more frames. The exact wording of the original frame was generally used, and the meaning was carefully maintained if any rewording or condensation was required. The time allowed for each frame was based on the mean times required by students who studied the teaching machine version on a self-paced basis. Figure 15 illustrates some visuals used for a programed sequence on television.

3) The visuals which were to be used by the live instructor were enclosed in acetate folders for grease pencil use and easy removal.
4) A detailed script for the instructor which included the audio commentary was prepared for use on the prompting device.
5) The final step was the preparation of a complete, detailed script for the director.

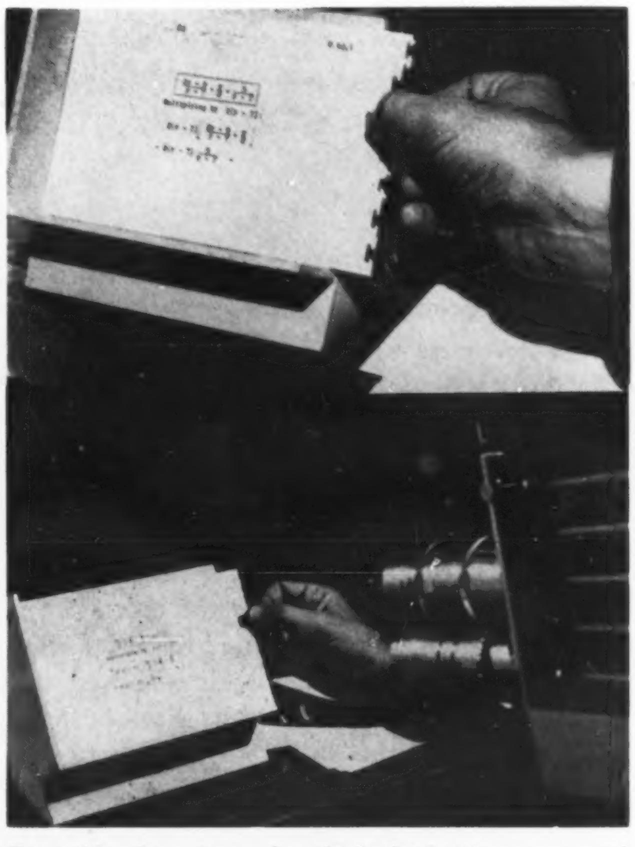


Figure 15. Two views of method of televising programed material in algebra course.

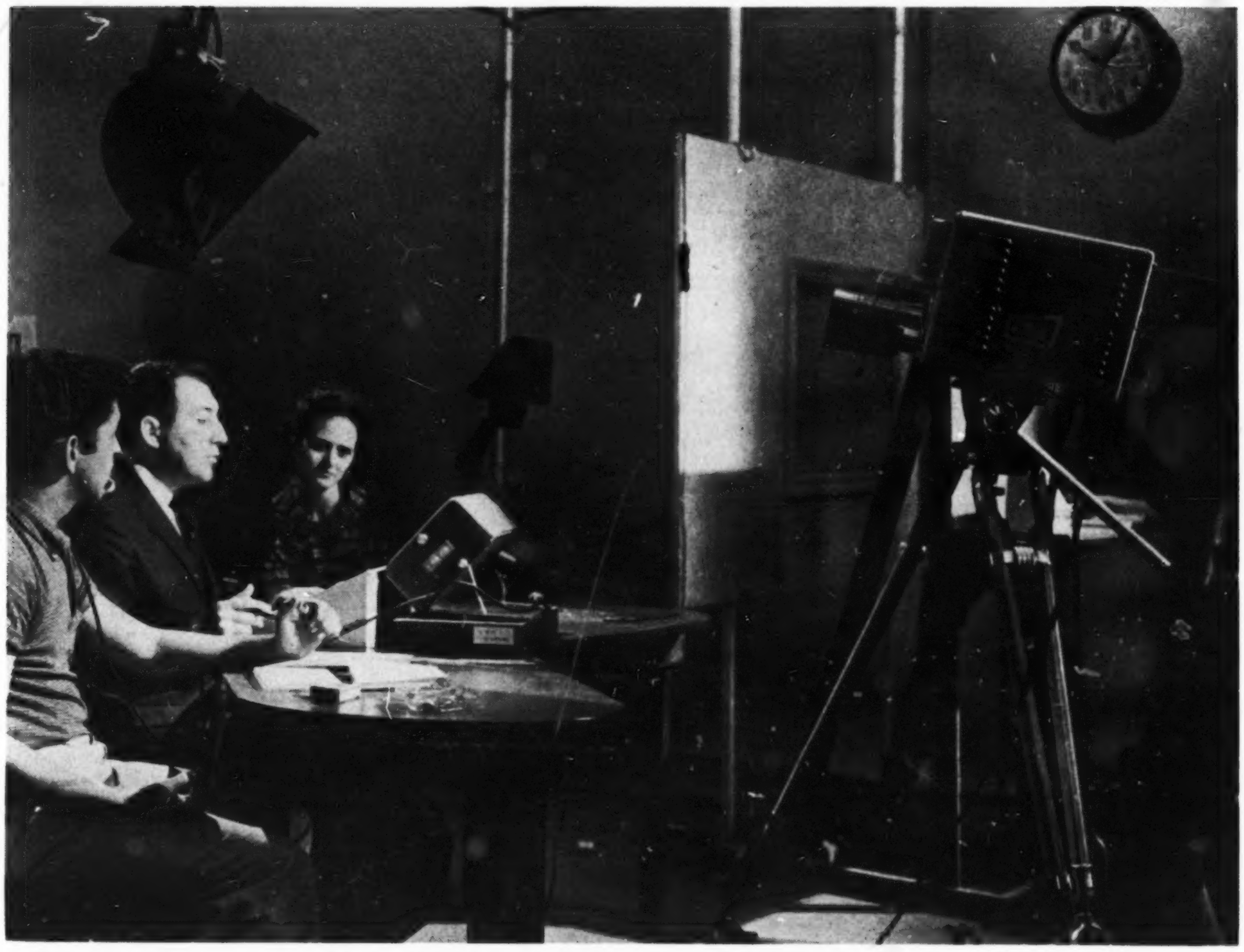


Figure 16. Studio arrangement during televised presentation of programed course.

Television Studio Procedures

Two cameras were used to televise the presentations. One was used to pick up the typewritten programed visuals. The other was used to televise the instructor.

Both cameras were of the studio Vidicon type with viewfinders and multiple lenses mounted on turrets. On one camera a four-inch lens was used to televise the typewritten cards containing the programed material. (See Figure 15.) The camera which televised the instructor was equipped with a zoom lens. Mounted above the instructor's desk was a mirror which reflected the visual material used by the instructor. Figure 16 shows the arrangement of the studio.

In operation this camera was focused on the instructor to give a head and shoulders view. When the instructor turned to his visual materials, the camera would be tilted up to the mirror and the lens "zoomed" to provide an image of the correct size. In order that writing and printing would appear with the correct orientation, this camera was operated with "reverse sweep" so that the mirror image was correctly presented. Since the instructor's hands were not shown in the direct view of him, the audience was not aware of the fact that the image had been reversed.

Such an arrangement proved to be very convenient since the instructor could easily go from lecture to visual materials without moving from his desk. Also, because of the need for a precise presentation of the material, the television prompting device was used. The display box for this equipment was placed in front of the instructor's lectern so that he could follow his script with ease.

Seated to one side of the television instructor was an assistant who handled the time cues for controlling the presentation of each frame. This individual had a list of the frames with a pre-determined time cue for each. He was also provided with a clock. When the appropriate time for a given frame had elapsed, this assistant would push a button which actuated an audible chime. This served as a signal for the camera operator to change the card to the next frame. It also alerted the viewing students to a change of frame.

The same assistant, who was the general coordinator for each television presentation, also cued the instructor when he was to make an audio comment; he also cued the director in the control room (by means of an intercom headset) when the scene was to shift from the program to the instructor or vice versa.

On the other side of the instructor was a second assistant who operated the television prompting device.

It might be thought that such a procedure was rather complicated and involved too many persons (the instructor, two assistants, two camera operators, and the director). In actual practice it worked very smoothly, and gave great flexibility. However, if such a method of presenting programed material were to be adopted for purposes other than research and development, it would undoubtedly be desirable to record the presentations on film or video tape for repeated use, after the format had been adequately developed and tested.

EFFECTIVENESS OF TELEVISED PRESENTATIONS

Since the results of Experiments I and II (discussed in earlier sections) supported the procedure of using programed materials for group instruction under externally-paced conditions, and since techniques had been developed for modifying and adapting programed mathematics material for televised presentation, the next step was to conduct an experimental assessment of the teaching effectiveness of the "programed" television version.

Consequently, the televised presentation of the adapted programed mathematics course (Units 1-28 of the previously described Contemporary Algebra program) was compared with an individual self-paced teaching machine presentation of the same original programed materials. The experimental comparisons were made on the criterion of achievement as measured by unit-tests, midterm, and final examinations.

METHOD

Subjects

The subjects were 63 students enrolled for credit in Mathematics 2 for the 1962 Spring Term at The Pennsylvania State University.

Materials

The first 28 units of the Contemporary Algebra program were used in this experiment. Each unit was supplemented by a study guide (previously described) which summarized the content of that unit and provided problems for homework assignment.

The sum of the scores on the unit-tests (each comprising from 11 to 17 multiple-choice items) described earlier was used as a criterion measure. Sixty-item comprehensive midterm and final examinations were also used.

The *School and College Ability Test* (SCAT)[1] was used as a pre-treatment measure of verbal and quantitative ability.

Treatments

1) *Televised Presentation of Programed Materials.* Programed units 1-28 were adapted

---

[1] Educational Testing Service, Princeton, N. J.

for televised presentation and used as the stimulus material for the television version. The subjects in this treatment proceeded through the program as a group, on a controlled-pace basis. Overt, constructed responses were written on prepared answer sheets similar to those used for the treatments in the previous two experiments.

The units were studied at the rate of one unit per class period. Each unit required not more than approximately 65 minutes (overall time); this left about 10 minutes of each 75-minute class session for the administration of the unit-test.

2) *Teaching Machine Presentation*. The programed materials were presented via teaching machines in a manner identical with the procedure described in Experiment I.

Students progressed through the programed units on a self-paced, time-limited (75 minutes) basis. Slow-working students were permitted to begin early as was allowed in Experiment I.

Experimental Design

The available population of 63 subjects was randomly assigned to the two treatment groups.

All subjects were pre-tested on the verbal and quantitative sections of SCAT; these were intended for use as adjusting variables in the statistical evaluation of the outcome of the experiment.

The treatments were imposed over one complete term of 10 weeks, under course-credit conditions.

A unit-test was administered immediately upon the completion of each unit. The unit-test scores were posted for the information of students at the beginning of the next class session.

Study guides were distributed at the termination of each unit (following the unit-test) and appropriate problems were assigned. All individual assignments were uniformly graded and returned to the students at the next class meeting.

The midterm examination was administered after the fifteenth unit and the final examination was administered after the twenty-eighth unit.

Except for differences in method of presentation, freedom for each "machine" student to progress at his own rate, and the option for slow-working "machine" subjects to begin early, administrative procedures were uniform over both treatment groups.

Comparisons were conducted on each of the three criterion measures:[1]

1) summed unit-test scores

2) midterm examination scores

3) final examination scores

Analysis of covariance was applied where the requirements for application of that model were satisfied; analysis of variance was used in one comparison where the covariance assumptions could not be met.

RESULTS

Analysis Based on the Summed Unit-Test Scores

An analysis of covariance was conducted to evaluate the difference between adjusted sample means. Examination of the verbal and quantitative SCAT scores as prediction of the achievement criterion (summed unit-test score) revealed that the quantitative section alone predicted for all practical purposes, as well as the weighted combination of verbal and quantitative sections:, therefore, the quantitative section of SCAT was used as a single adjusting variable.

The routine tests for significance of regression, homogeneity of regressions, and homogeneity of variances were conducted to ascertain whether requirements for application of the analysis of covariance model were satisfied. Results of these tests indicated that the basic assumptions had been met, and consequently the covariance analysis was carried out. The summary data for this analysis are given in Table 16.

---

[1] Kuder-Richardson Formula 20 reliability for the midterm exam, r = .70; for the final exam, r = .81. The reliability of the summed unit-tests totaling 410 items was .94.

TABLE 16

ANALYSIS OF COVARIANCE FOR SUMMED UNIT-TEST SCORES

Component	S.S.	df	V	F	p
Treatment	1028.62	1	1028.62	1.40	>.05
Error	43977.48	60	732.96		
Total	45006.10	61			

The F-ratio of 1.40 indicated that the differences among means were not significant at the .05 level. Therefore, in this experiment, there was no significant difference in the teaching effectiveness between the TV and teaching machine presentations, when compared on summed unit-test scores.[1]

The summary of observed and adjusted means is presented in Table 17.

TABLE 17

OBSERVED AND ADJUSTED MEANS FOR SUMMED UNIT-TESTS

Treatment	N	SCAT Quantitative Test Mean	Summed Unit-Test Observed Mean	Summed Unit-Test Adjusted Mean
Teaching Machines	36	37.19	330.39	331.09
Television	27	37.96	339.93	339.00

Analysis Based on Midterm Test Scores

It was the intent initially to apply the analysis of covariance model in order to compare treatment groups with respect to the midterm test scores. However, the test for homogeneity of regressions revealed that this requirement was not satisfied; therefore, the experimental comparison was made by applying analysis of variance, since all requirements for application of this model had been met.

The obtained F-ratio of 2.134 (1,61 degrees of freedom) indicated that the differences between the means were not significant at the .05 level. In this experiment, then, the TV and teaching machine presentations were equally effective with respect to achievement as measured by the midterm examination.

The summary data for the analysis of variance appear in Table 18.

The observed means are given in Table 19.

TABLE 18

ANALYSIS OF VARIANCE FOR MIDTERM TESTS

Component	S.S.	df	V	F	p
Treatment	88.7314	1	88.7314	2.134	>.05
Error	2536.1575	61	41.57635		
Total	2634.8889	62			

TABLE 19

OBSERVED MEANS FOR MIDTERM TESTS

Treatment	N	Midterm Observed Means
Teaching Machines	36	46.75
Television	27	49.15

Analysis Based on Final Examination Scores

Analysis of covariance was used to evaluate the outcome of the experiment with respect to final examination scores.

Prior to application of the covariance model, the routine tests were conducted in order to insure satisfaction of the requirements of that model. Since the hypotheses of homogeneous regressions and homogeneous variances were accepted, and there was a significant regression, the analysis of covariance method was employed.

The obtained F-ratio of .21 (1,60 degrees of freedom) indicated that the differences among means were not significant at the .05

level. Consequently, it was concluded that, with respect to achievement as measured by final examination scores, the TV and teaching machine presentations of the programed materials did not differ in effectiveness.

The summary data for the analysis of covariance are given in Table 20.

TABLE 20

ANALYSIS OF COVARIANCE FOR FINAL EXAMINATION SCORES

Component	S.S.	df	V	F	p
Treatment	6.84	1	6.84	.21	>.05
Error	1965.80	60	32.76		
Total	1972.64	61			

The summary of observed and adjusted means is given in Table 21.

TABLE 21

OBSERVED AND ADJUSTED MEANS FOR FINAL EXAMINATION

Treatment	N	SCAT Quantitative Test Mean	Final Test Observed Mean	Final Test Adjusted Mean
Teaching Machines	36	37.19	47.36	47.45
Television	27	37.96	48.37	48.17

DISCUSSION OF RESULTS AND CONCLUSIONS

Part I: Adaptation of Program to Television

Having earlier ascertained that the externally-paced presentation of programed materials to groups was a promising undertaking, pilot steps were initiated for the purpose of developing techniques for modifying and adapting the mathematics program for televised presentation.

Major questions centered around the optimal use of the auditory channel, rate of presentation, where and how to augment the program with a live instructor, and how to produce and televise the visual materials.

The question of the rate of presentation was answered by Experiments I and II, and the questions of use of the auditory channel and instructor appearances were answered through preparation and presentation of pilot lessons.

Through several trial-and-error developmental stages, routine procedures were established for modifying and adapting programed materials for televised presentation.

Part II: Experimental Results

The television and teaching machine presentations of the programed course were compared on three criteria:

1) summed unit-tests
2) midterm examination
3) final examination

Although the hypotheses which were tested were not independent, it was judged that the opportunities for gaining additional information warranted the carrying out of an evaluation on each of the three criteria. The reader is cautioned to take cognizance of the fact that the criterion tests were substantially correlated.

Since the hypothesis of equal means could not be rejected for each of the three comparisons, the conclusion must be reached that, under the experimental conditions imposed, the televised and teaching machine presentations of the program materials were equally effective.

This evidence supports that of the preceding experiments (I and II), which showed that students have rather wide tolerance limits for variations of pace.

Of primary importance, however, is the conclusion that the presentation of programed materials over television can be an effective and feasible procedure.

This finding suggests a number of possibilities for the wide distribution of programed materials by the use of television, both closed-circuit and broadcast, and of integrating programed sequences with lecture-demonstrations over television. The programed sequences would provide for practice and application of principles, for overt responses on the part of all learners, and for immediate reinforcement through provision of knowledge of results. More generally, programed televised instruction helps to insure *persistent, intensive*, and *appropriate* responses by students

EXPERIMENT IV: Paired vs. Individual Study of Programed Instruction in Contemporary Algebra

by Walter Dick[1]

BACKGROUND AND PURPOSE

Consideration must be given to one of the most concrete objections to programed instruction which is expressed both by students who have used these materials and by the public. This objection can be stated as the lack of the usual type of interaction between teacher and student. The layman is often skeptical of a teaching method (or more precisely, a learning method) which permits no such interactions. In some studies (Gagné and Dick, 1962, for example)[2] students have indicated that the day-in, day-out use of programed learning is monotonous and tiring, and earlier experimentation on this project confirmed this finding. Although the material itself may be interesting, the problems of motivation require study.

Pressey (1959) has expressed his concern for this isolation of the learner. He states that "... programing would seem to emerge as of great promise especially for certain kinds of materials and certain purposes, but needing coordination with research on student needs, knowledge, and errors. And there may be danger that learning be made too much something one attempts in a booth alone with a machine and enmeshing programed matter; most effective learning may rather be a varied, lively, challenging experience with frequent interplay of minds."[1]

In order to attempt to meet some of the criticisms of programed instruction, to explore other aspects of the pacing problem, and to study programing for groups, an experiment was carried out to investigate the effect of pairing students for study of the program on Contemporary Algebra. Each student in the experimental group was paired with another student, and they worked together through the entire course. Incorrect responses were discussed, and the pace of the pair was determined by their combined rates and need for discussion. The performance of the paired students was compared with that of students who worked alone. The pairing procedure also provided a student with the opportunity to ask questions of his partner about material with which he was unfamiliar, and this presumably removed him from the "social isolation" which characterized earlier research procedures with programed instruction.

It may also be hypothesized that the pairing procedure may serve as a substitute for the student-teacher or student-proctor relationship which is absent in the "teacherless" use of a program. The students' interaction could serve to reinforce socially the behavior of each of the members, therefore, increasing the perceived acceptance of the learning situation.

There were two other areas of interest with which this study was concerned. One was the question of predicting final performance: what is the relationship of (1) mathematical ability, and (2) verbal ability to the learners' performance on the program's evaluation measures? Since a mathematics program was employed in this study, it could be assumed that mathematical ability would be the major factor influencing student performance. However, self-instructional programs have been characterized as being highly verbal. This suggests that verbal ability itself might be the dominant factor contributing to final performance.

---

[1] Dick, Walter, *Effects of Paired Student Interaction with Programed Instruction.* Unpublished Master's Thesis, The Pennsylvania State University, 1963.

[2] Gagné, R. M., and Dick, W., "Learning Measures in a Self-instructional Program in Solving Equations," *Psychol. Rep.*, 1962, 10, 131-146.

[1] Pressey, S. L., "Certain Major Psycho-educational Issues Appearing in the Conference on Teaching Machines." In *Automated teaching: the State of the Art* (ed. Galanter), Wiley, New York, 1959, p. 193.

The third area of interest in this study was that of student attitude under conditions involving the study of a substantial amount of programed material. Most of the studies of student attitude with programed instruction have been conducted with programs of 50 to 100 frames (Feldhusen, 1961).[1]
This seems to be an insufficient exposure to a novel learning technique to justify any generalizations concerning attitudes toward the use of programed instruction in general. Since the programed material (28 units) employed in the present study contained 3,512 frames, it was believed that a more accurate measure of student attitudes could be obtained. After working approximately 30 hours on the program, the students should have either confirmed or discredited any doubts or hopes which they developed during their early exposure to the program and should thereby reflect a more stable, informed attitude toward programed instruction.

METHOD

Subjects

Seventy students, who were enrolled in Mathematics 2 at The Pennsylvania State University during the Winter Term of 1962, served as subjects.

Materials

The first 28 units of the Contemporary Algebra program, previously described, were used in this experiment. Each unit had an 11 to 17 item multiple-choice test (also previously described), a study guide which contained a review of important concepts from the unit, further discussion, "solved" problems, and problems for assignment. The 60-item cumulative midterm and final examinations and a 25 item "transfer" test were also used. The transfer test contained items designed to test for concept understanding rather than specific knowledge.

Educational Testing Service's *School and College Ability Test*[1] was used as a pretreatment measure of verbal and quantitative ability. Hartman's[2] course attitude questionnaire, described under Experiment II, was used to evaluate student attitude toward the course and the program.

Treatments

The sample population of 70 students was randomly assigned to two groups which met in adjacent rooms. During the first meeting both groups were given a general orientation to the course. Both sections were given the same general instructions, with the exception of an explanation of the pairing procedure which was to be used by the experimental group. The "teaching machine" version of the program was used by both groups. Instead of using the aluminum "envelope," however, the students used card masks which could be moved down each page of the program, successively revealing a frame, then a correct answer.

Individual Group Treatment. The students who worked as individuals noted their "starting time" on their answer sheets and then proceeded through the program unit at their own rate by sliding a mask down the page, successively revealing a frame and then the answer. Upon completion of a unit, they took the unit-test and received a study guide and study assignment for the next session. At no time during the class were the students permitted to ask questions of the proctor except in the case of misprints in the program.

Paired Group Treatment. The students who worked in pairs were told that they should sit side-by-side with the program placed between them; that they should read the first frame and write their answers on their

---

[1] Feldhusen, J. A., "Reactions of College Students to a Self-instructional Teaching Device and Programed Instruction," *AID*, 1961, 1, 37-38.

[1] Educational Testing Service, Princeton, N. J., Kuder-Richardson Formula 20 reliability, r = .859.

[2] Hartman, *op. cit.*, p. 26.

own individual answer sheets. When both students were finished with a frame, the mask was moved down the page to reveal the correct answer. If both students were correct, they were to go on to the next frame. If one or both students were wrong in their answer, they were instructed to check over the item and discuss the material in the frame until both members were satisfied that they understood it. The procedure for tests and study guides was identical to that of the control group. Figure 17 shows the classroom arrangement for the individual and paired group treatments.

Experimental Design

This experiment employed a single independent variable, namely the pairing of students using programed instruction. All the students were randomly assigned to either the control (individual group) or the experimental (paired group) treatments. The students in the experimental group were then *randomly assigned* to pairs. The experiment was conducted over the 10-week Winter Term of 1962 during the three regularly assigned 75-minute class periods per week.

The study of the program units was initiated during the first class following the orientation session and continued through the 28 units; the SCAT test was administered at a special session preceding the administration of the units. The only interruption of this schedule was for the administration of the midterm examination after the fifteenth program unit. The regular study guide for the twenty-seventh unit was replaced by the attitude questionnaire, which was completed at home by the students and returned at the next class meeting. The final examination and the transfer test were administered following the twenty-eighth unit.

The learning sequence of each student was controlled by the requirement that no new unit could be started until the preceding one had been completed. When a student (or a pair of students) missed a class, he (or the pair) attended one of the nightly make-up sessions in order to cover the missed unit. These make-up sessions were conducted in the same manner as the regular classes.

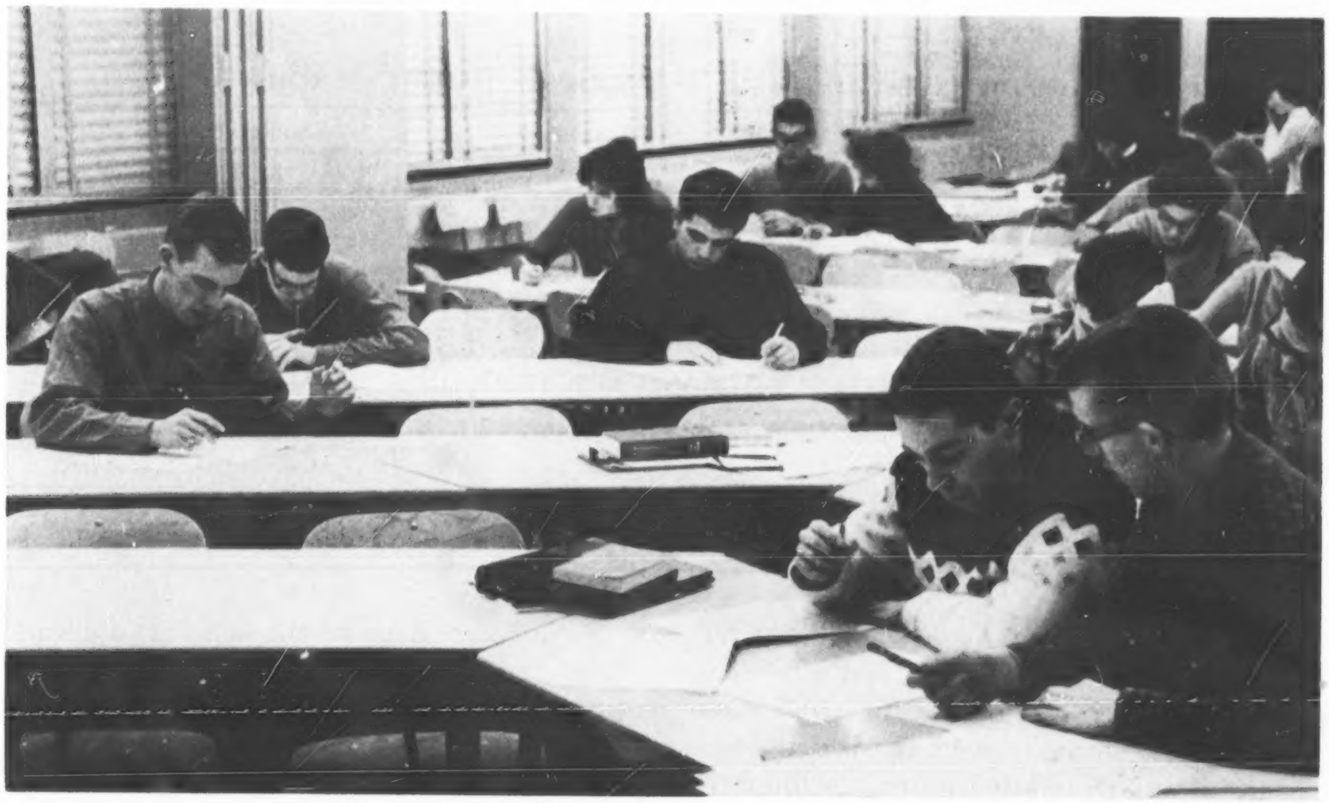


Figure 17. Classroom arrangement showing individual and paired group treatments.

RESULTS

Midterm and Final Examination

Midterm Examination. On the midterm examination the paired group had a mean score of 49.3 out of a possible 60 points. The individual group had an average of 47.5. Having first ascertained that the requirements of significant regression, homogeneity of regressions, and homogeneity of variances were satisfied, analysis of covariance was applied to evaluate the difference between adjusted means. The SCAT test was used as the adjusting variable. The obtained F-ratio of .91 (1,67 degrees of freedom) indicated that the differences among means were not significant at the .05 level. A summary of the covariance analysis for the midterm examination appears in Table 22. The observed and adjusted means are given in Table 23.

TABLE 22

ANALYSIS OF COVARIANCE
FOR THE MIDTERM EXAMINATION

Component	S.S.	df	V	F	p
Treatments	25	1	25	.91	$p>.05$
Error	1847	67	27.6		
Total	1872	68			

TABLE 23

PREDICTOR, OBSERVED MIDTERM, AND ADJUSTED MIDTERM EXAM MEANS

Treatment	N	Total SCAT Mean	Midterm Exam Obs. Mean	Midterm Exam Adj. Mean
Paired Group	36	79.0	49.3	49.0
Individual Group	34	75.4	47.5	47.8

Final Examination. Again using the SCAT as the predicting variable and first ascertaining that the requirements for the application of the covariance model were satisfied, the difference between adjusted means on the final examination was evaluated. The obtained F-ratio of .01 (1,67 degrees of freedom) indicated that the differences among means were not significant at the .05 level. A summary of the analysis of covariance appears in Table 24, with the observed and adjusted means given in Table 25.

TABLE 24

ANALYSIS OF COVARIANCE
FOR THE FINAL EXAMINATION

Component	S.S.	df	V	F	p
Treatments	.5	1	.5	.01	$p>.05$
Error	2563.6	67	38.3		
Total	2564.1	68			

TABLE 25

PREDICTOR, OBSERVED FINAL, AND ADJUSTED FINAL EXAM MEANS

Treatment	N	Total SCAT Mean	Final Exam Obs. Mean	Final Exam Adj. Mean
Paired Group	36	79.0	48.4	48.0
Individual Group	34	75.4	47.5	47.9

Transfer Test

The test of transfer was composed of 25 multiple-choice items designed to test for concept attainment rather than specific knowledge.[1] On this test the mean performance of the paired group was 14.0 and for the individual group 12.6. The difference between adjusted means yielded an F-ratio, $F = 1.79$ (1,67 degrees of freedom) indicating that the difference between the performance of the groups was not significant at the .05 level.

Summary data for the analysis of covariance and observed and adjusted means are presented in Tables 26 and 27, respectively.

---

[1] The Kuder-Richardson Formula 20 test of reliability was .54.

TABLE 26

ANALYSIS OF COVARIANCE
FOR THE TRANSFER TEST

Component	S.S.	df	V	F	p
Treatments	22	1	22	1.79	p>.05
Error	821	67	12.3		
Total	843	68			

TABLE 28

ANALYSIS OF COVARIANCE
FOR THE SUMMED-UNIT TESTS

Component	S.S.	df	V	F	p
Treatment	197	1	197	.23	p>.05
Error	58388	67	871		
Total	58585	68			

TABLE 27

PREDICTOR, OBSERVED TRANSFER, AND ADJUSTED TRANSFER TEST MEANS

Treatment	N	Total SCAT Mean	Transfer Test	
			Obs. Mean	Adj. Mean
Paired Group	36	79.0	14.0	13.8
Individual Group	34	75.4	12.6	12.8

TABLE 29

PREDICTOR, OBSERVED SUMMED-UNIT, AND ADJUSTED SUMMED-UNIT TEST MEANS

Treatment	N	Total SCAT Mean	Summed-Unit Tests	
			Obs. Mean	Adj. Mean
Paired Group	36	79.0	335.9	333.3
Individual Group	34	75.4	327.3	330.2

The raw scores of the daily unit-tests were summed over all 28 units.[1] The total number of test items was 410. The paired group had an average of 335.9 correct items while the individual group averaged 327.3. When converted to percentages, the averages are 81.9 and 79.8 per cent, respectively. Analysis of covariance was again employed to test for significance of difference between adjusted means of the total scores for all 28 daily tests. This procedure is appropriate when the daily quizzes are considered as one large test consisting of 410 items administered over 10 weeks. The obtained F-ratio of .23 (1,67 degrees of freedom) indicated that the differences between means were not significant at the .05 level.

The summary data for the analysis of covariance are presented in Table 28; observed and adjusted means appear in Table 29.

Group performances on each unit-test are displayed graphically in Figure 18.

Results of Analysis of Learning Time

The students' starting and stopping times were combined to give an average time on each unit for each group. The average unit times for the paired groups ranged from 45.4 to 73.7 minutes. The range for the individual group was from 45.5 to 70.4 minutes. The mean unit time for the paired group was 62.2 minutes; that for the individual group was 58.5 minutes, a difference of 3.7 minutes. The average time for the entire program was 29.0 hours for the paired group and 27.3 hours for the individual group. This was a statistically significant difference, $t = 2.76$, $p < .01$. However, the difference is of little practical significance.

Prediction Results

In order to investigate the role of verbal and quantitative ability in performance on the program, multiple correlation and regression procedures were used on the results of the final examination. The multiple $R = .432$ for the paired group; $R = .392$ for the individual group. However, since these R's were computed for small samples ($N < 50$) a correction for bias was necessary. The corrected R's for the paired and individual groups were .402 and .358, respectively. When the corrected multiple correlations are squared, the resulting figures indicate that 16.2 per cent of the

---

[1] The reliability of the summed unit-tests totaling 410 items was .94.

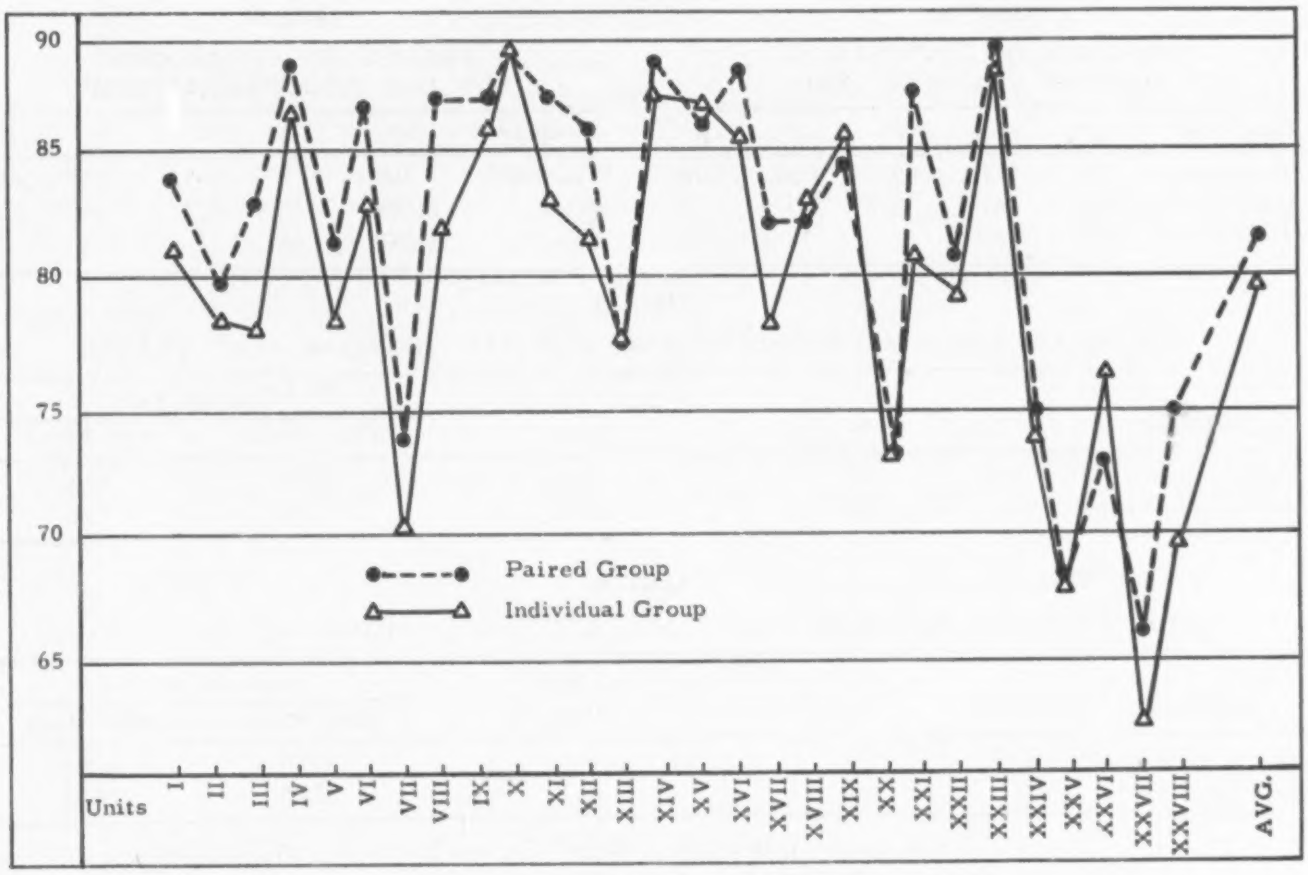


Figure 18. Performance on the 28 Daily Quizzes Expressed in Per Cent of Maximum Possible Score.

variance on the final examination for the paired group and 12.8 per cent for the individual group are accounted for by verbal and quantitative ability.

For the paired group uncorrected R^2 equaled .186, which was composed of a quantitative factor of .072 and a verbal factor of .114. It was determined that 38.7 per cent of the "accounted-for" variance is represented by quantitative ability and 61.3 per cent by verbal ability.

Using the same procedure for the individual group, uncorrected R^2 equaled .153, with a quantitative factor of .118 and a verbal factor of .035. These figures indicated that 77.1 per cent of the accounted-for variance is represented by quantitative ability and 22.9 per cent is represented by verbal ability.

Attitude Questionnaire

All of the items on the attitude questionnaire were scored on a one- to five-point scale. Strong agreement with a positive statement about the course was given a score of 1, indifferent attitude was scored a 3, and a strongly negative response was scored a 5.

The first section of the questionnaire attempted to measure the general attitude of the students toward the course. The average score on this section for the paired group was 3.30 and 2.99 for the individual group, indicating a neutral attitude. The difference between the groups was not significant at the .05 level.

In the second section of the questionnaire, as constructed by Hartman,[1] the students were asked to evaluate the teacher of the course. For purposes of this experiment, these questions were re-worded to evaluate the programed course rather than the teacher.

The overall attitude toward the program itself, exclusive of the other components of the course, was 2.71 for the individual group; 3.01 for the paired group. This difference was not significant. The averages represent a wide range of attitudes varying in both groups from highly favorable to highly unfavorable.

---

[1] Hartman, *op. cit.*, p. 26.

Product-moment correlations were computed between (1) general course attitude and performance on the final examination, and (2) attitude toward method of instruction with performance on the final examination. For the individual group r = -.28 for course attitude and -.25 for program attitude. These correlations are not significant. The correlation between general course attitude and the final examination for the paired group was r = -.02, which implies no relationship. The relation between program attitude and final examination for the pairs was r = .12, which is also non-significant.

General Reactions

The students were also requested to make general comments about the entire course. Of the 69 students who completed the course, 26 out of 35 in the paired group and 24 out of 34 in the individual group offered additional comments. In the following paragraph the number of students making particular comments is noted in parentheses. In reference to the entire course, there was little agreement in the remarks from the students in the paired group. They noted that the units studied in the last week and a half were too difficult (2); first period was too early in the morning (2); dislike of the restriction on asking of specific questions of the proctors during the class period (1); and dislike of the feeling of just being in an experiment (1). In contrast, all of the general comments (8) from the students in the individual group were related to the idea that the course became a "monotonous routine," that it was "impersonal" and "tiring." The students "missed asking questions" and "developing a speaking knowledge of math" and even noted "the inability to get to know one's own classmates." No comments of this nature were made by members of the paired group.

DISCUSSION OF RESULTS AND CONCLUSIONS

Individual vs. Paired Learning

The results of the group performances on the midterm and final examinations, the test of transfer, and the total daily unit-test points indicated no significant differences at the .05 level. This result indicates that the paired group performed at least as well as the conventional type of individual group despite the fact that—because the students were randomly assigned to pairs—no consideration was given to the background, personality, or ability of the students.

A significant difference was obtained on the average unit times of the two groups. The paired group required an average of 3.7 minutes more per unit than the individual group. Although this result is statistically significant, it is of little practical significance when the total range of the learning times is considered. This difference in learning time may partially be attributed to the random assignment of pairs. For example, in the pairs in which there was a large difference in reading speed, one member often simply sat and waited for his partner to complete a frame and this would, of course, increase the total group learning time.

Prediction of Performance

The results of the multiple correlation and regression analysis indicate the relative importance of verbal and quantitative abilities in the prediction of performance. It should be noted, however, that the total variance accounted for by combined verbal and quantitative ability is quite small, less than 20 per cent. The reader must be further cautioned on the generalization of these interpretations since the small N and absence of cross validation raise the question of whether these results are reproducible.

It was shown that for the individual group about 77 per cent of the *accounted-for* variance represented quantitative ability, and approximately 23 per cent represented verbal ability. These figures suggest that a student who works alone and has a great deal of learned quantitative ability may perform well on this algebra program, while a student who lacks this ability may perform poorly. Verbal ability appears to play a rather insignificant role for the individual who works alone.

However, when a similar analysis is performed for the paired group, only about 39 per cent of the accounted-for variance is represented by quantitative ability, while about 61 per cent is represented by verbal ability. Seemingly, when students work in pairs and discuss the program, verbal ability becomes a more important factor, a factor which may help a student with a low quantitative ability score and a high verbal score to perform better than if he worked alone.

Student Attitudes Toward Programed Instruction

The results of the attitude questionnaire completed by all the students following the twenty-seventh unit were somewhat contrary to previously reported results, i.e., those of Feldhusen (1961).[1] The students indicated that they were not overly enthusiastic about programed instruction as it was employed in this experiment. The average ratings for the two groups indicate that the students were relatively neutral toward the material.

The results of the general comments about the course further support the idea that the impersonal conditions and lack of interaction in individual study make this an undesirable way to use programed materials. Many of the students in the individual group noted that they disliked the monotonous routine which was imposed by their experimental treatment. While the members of the paired group did not complain of the monotonous routine (suggesting that the interaction of the students may have served as a substitute for a teacher or proctor relationship), many did report that they felt that in the future students should be paired on the basis of *equal ability*. Observation of the pairs which seemed to have the greatest difference in "working speed" seemed to indicate that delays were often caused, not by the inability of one of the students to respond to the material, but rather by the slower reading and comprehension rates of the other student.

It would seem advisable to investigate systematically the effects of other methods of pairing students as an approach to analysis of student interactions relative to possible ways of enhancing the interest of programed materials.[1]

---

[1] Feldhusen, *op. cit.*, p. 40.

[1] A follow-up study involving the pairing of students to study the programed course in English grammar is described later in this report.

PART III

RESEARCH ON METHODS AND MEDIA FOR PRESENTING PROGRAMED ENGLISH GRAMMAR

Development of the English Program

by A. Funk and L. P. Greenhill

Introduction

Research on the programing of English grammar and the experiments involving the use of closed-circuit television for the presentation of the programed course closely paralleled the studies already reported for mathematics.

One of the main objectives was to select a distinctly different subject matter area and to apply similar procedures in order to test the generality of the results of programing. All too frequently educational experiments are not replicated or are not repeated in more than one subject area, and this seriously limits the degree to which the findings can be generalized to other situations, contents, or students.

This part of the project was conducted as a cooperative enterprise between the Division of Academic Research and Services[1] and the Department of English at The Pennsylvania State University.

Philosophy in the Teaching of English Grammar

The approach to the teaching of English grammar in the public schools and colleges has generally been "traditional." A traditional approach may be described as a prescriptive method; that is, based on rules, whereby English sentence structure is analyzed according to the structure of Latin grammar. Many teachers of English have felt that such a point of view was unrealistic since there is no such thing as a universal grammar; each language has its own peculiar structure and it is therefore illogical to assume that one could impose Latin rules of grammar and the Johnsonian grammar of eighteenth century England on current American English structures. Secondly, a prescriptive method assumes that a language is a static thing. Latin, no longer in use as an active language, can be fitted into this category, but English is a growing and expanding language. Traditionalists may have found "beauty" in the static structure of Latin as a referent for what is "proper" in English; but since English is a changing language, the application is not realistic. Thus, traditionalists must, of necessity, teach a pattern that is disparate from actuality.

For centuries men have been studying the structure of language, but not until recently have the findings of the structural linguists reached the classroom. The structural linguistic approach, as opposed to the traditional approach to teaching grammar, rests on these premises:

1) Writing, being a symbolization, is secondary to speech, though not necessarily less important.

2) Correctness in speech or writing is relative to time, place, and circumstance.

3) Change is a normal and necessary characteristic of language, though not necessarily of writing.

4) There is no such thing as a universal grammar; each language has its own peculiar structure.

5) Any normal child of five or six is a perfect speaker of his language in the sense that he can manage its sounds and syntax automatically.[1]

---

[1] Subsequently renamed "University Division of Instructional Services."

[1] Roberts, Paul, *Understanding English*, Harper and Brothers, New York, 1958, pp. x-xi.

A linguistic approach is therefore descriptive of the existing changing language, rather than prescriptive, i.e., based on attempts to pigeon-hole American English constructions in the structure of Latin grammar.

In planning and writing the program for use in the present studies, it was decided that an approach would be used that was based more on linguistic principles and less on the traditional method of teaching English grammar.

The term "linguistics" as used in this report means a systematic description of the structure and operational procedures of the English language. For example, word order in English is rigid. Form classes (parts of speech) are marked by derivational affixes and function words, such as determiners (which mark nouns) and auxiliaries (which mark verbs). Thus, both the form and function of words are keys to the structure of the language and are more accurate than definitions.

Traditional grammar teaches, by definition, that a noun is the name of a person, place, thing, or idea. By this definition we would have to say that *sky* is a noun in the sentence, "The sky is blue," because it names a thing. Similarly we would have to say that *blue* is a noun because it is the name of a color. If we accept the traditional definition that one of the functions of an adjective is to name a quality, we would have to identify *blue* as an adjective. As a result of this definition, we would have to say in the sentence, "His bravery was unusual," that *bravery* is an adjective because it names a quality. Hence, it can be seen that definition becomes an inadequate and inaccurate means of analyzing language.

A linguist describes nouns as belonging to a form class. That is, nouns can be identified by a set of forms which they have in common. For example, because word order is rigid in English, only nouns will fill the blank in "He found his ...." or "It was filled with ...." The suffixes *-ness, -tion,* for instance, are another characteristic common to nouns. A recognition of word order and characteristics common to a given word class is a more accurate method of describing the characteristics of an existing language than is a complicated set of conflicting rules.

These principles were kept in mind when planning and writing the programed course in English grammar which was used in the experiments reported in the following chapters.

Purpose of the Course

The programed course was designed to be used as an adjunct to the freshman course in English composition at The Pennsylvania State University. All incoming freshmen are required to take an English Placement Test. Those scoring below a specified cut-off point were formerly required to study and pass a non-credit course in English usage (English 0) before they could be admitted to the freshman composition course.

This procedure has been changed, and under the current plan, students scoring below the cut-off score of 45 on the English Placement Test are required to take supplementary instruction in English grammar (English RE) concurrently with the freshman English composition course. The English composition course (which includes written assignments) requires three 75-minute class periods a week for 10 weeks. The supplemental instruction in English grammar requires two additional 75-minute periods a week.

These students are otherwise qualified to undertake university work, but for one reason or another their previous training in English grammar has been inadequate. The programed course, which was developed for the present study, was intended to provide this supplemental instruction in English grammar. It was anticipated that such a programed course, if successful, would be useful not only for supplemental instruction at the university level, but would also be appropriate for use in secondary schools.

Selection of Topics for Inclusion in the Programed Course in English Grammar

Several members of the English Department at The Pennsylvania State University determined the content of the course.[1]

On the basis of discussions, the topics listed in Table 30 were selected for inclusion in the programed course.

TABLE 30

PROGRAM UNIT TITLES
ENGLISH GRAMMAR

UNIT	*TITLE*
Intro.	Introduction to Programed Instruction
I	Nouns and Pronouns
II	Verbs
III	Subject-Verb
IV	Adjectives and Determiners
V	Adverbs and Qualifiers
VI	Prepositions
VII	Sentence Patterns (1-4)
VIII	Sentence Patterns (5-8)
IX	Verbal Adjectives
X	Verbal Nouns and Adverbs
XI	S-Groups (Adjective and Noun)
XII	S-Groups (Adverbial)
XIII	Punctuation
10	Spelling Units

The program was to be written in such a way that the students were expected to complete a grammar unit, a short spelling unit, and a unit-test within a 75-minute class period. The entire course was to be completed within 13 class periods at the rate of two per week, thus leaving time for the administration of pre-tests, a midterm test, and an end-of-course test within a 10-week term.

Preparation of the Program

The program was written by a team of three experienced teachers of English.[2]

After the first writing of each of the programed units, they were read and responded to by each of a group of four undergraduates under the direct observation of one of the program writers.

The observer made a record of all items which were incorrectly answered; she also discussed the content of each unit with each of the students in an effort to determine the extent to which the students had learned the concepts that were presented and where the difficulties lay in the incorrectly answered items. All of this information was used as the basis for the first revision of the program.

After the first revision, the program was tested with a group of approximately 20 students enrolled in English RE during the Spring Term, 1961, at The Pennsylvania State University.

These students completed two units each week. A third class period was conducted by an instructor in a conventional face-to-face lecture situation. Data on student error rates on the programed units, both by item and unit, were collected. These data were analyzed and the information formed the basis of the second revision.

The final testing of the program took place during the Summer Term, 1961, with a group of approximately 45 freshmen who had not passed the English Placement Test. In addition, an electronic operations recorder, which provided a response station for each of the students, was employed to record the time required by each individual to respond to each item of the program, as well as the total time required to complete each unit. These data were studied and the final revision was then undertaken. The timing data were later used to adjust and pace the materials for use on closed-circuit television.

Two members of the English Department at The Pennsylvania State University[1] and James D. Gordon, Professor of English at the University of Pennsylvania, read the entire program and provided many helpful suggestions for its improvement. Wherever possible, their suggestions were incorporated.

The number of frames in the 13 grammar units ranged from 53 to 174, with an average of about 100. The total number of frames

---

[1] This group included Henry W. Sams, Head of the English Department; John S. Bowman, Professor of English Composition; Agnes R. McElwee, Professor of English Composition; James P. Stewart, Associate Professor of English Composition; and Ann Funk, Research Assistant. Dr. James D. Gordon of the University of Pennsylvania and Dr. Nelson W. Francis of Brown University (then at Franklin and Marshall College) also served as consultants in the early planning stages of the project.

[2] Ann Funk and Agnes McElwee wrote the units on grammar. James P. Stewart wrote the spelling units.

[1] Henry W. Sams, Head of the English Department, and John S. Bowman, Professor of English Composition.

2

was 1,383. The 10 spelling units contained 350 frames.

In the original preparation of the program, each frame was written on a 5" x 8" card with the correct answer shown on the back. These cards were bound into units with loose leaf binder rings to form "programed books" for purposes of testing them with students. Students' responses were written on special answer sheets rather than on the program itself. (See Figure 19.)

The program was linear, rather than branched, in that the information was presented in small, logical steps and the student responded to each problem or question. After he was informed of the correctness or incorrectness of his response, he proceeded to the next frame.

Programing techniques are supported by modern learning theory, which has shown that learning will be accelerated if students make frequent responses and if correct responses are immediately rewarded or reinforced by knowledge of results.

Figure 19. Answer sheets used in the "programed" presentations.

Repetition is essential to ensure learning and retention of what is learned. For these reasons the program provided sufficient illustrative frames and built-in repetition. To fulfill the requirements of the project, the programed materials in card form were reworked (1) so that they could be used in a teaching machine, and (2) so that they could be presented over closed-circuit television in order to compare the effectiveness of the various media for presenting programed materials.

The Development of Ancillary Materials

Study guides were prepared—one to accompany each of the 13 units. These guides summarized the main topics in each unit and served in lieu of "lecture notes." The students were also given "take-home" tests which provided additional practice with the concepts developed in each unit.

Conversion to Teaching Machine Format

The programed units in card form were reproduced onto legal-size paper for use in the teaching machines which had been especially designed for this project. (See Figure 20.) On each sheet of paper a frame would appear, followed by the correct answer. This would be followed by another frame and correct answer until the sheet was full. The program would then continue on the next sheet.

The students successively exposed a frame, then an answer, then another frame by gradually withdrawing each page from the teaching machine so as to expose in the window the frames followed by reinforcement cues or answers.

The students' responses were written on specially prepared sheets which had answer spaces numbered to correspond with the program frames.

Adaptation for Televised Presentation

During the late summer and early fall of 1961 the English program was adapted for presentation over closed-circuit television. An attempt was made to keep the content of the television version as close as possible to the teaching machine version. However, certain changes were made in the television version. First, it was found that some of the frames in the original program were too long to be presented at one time on the television screen and still be legible when viewed from a distance of 12 screen widths from the screen (the maximum viewing distance). Consequently, some frames in the original version were divided into two frames for presentation over television. Second, since the audio channel is available in television, some material that appeared in printed form in the teaching machine version of the program was presented in the form of the spoken word over television.

The audio information consisted either of explanatory or prefatory comments concerning something which appeared in print on the television screen, or it sometimes repeated what was on the screen for purposes of emphasis. Sometimes the correct answers were given in printed form, sometimes in spoken form, and occasionally in both forms.

For the initial try-out of the television version in the fall of 1961, the narrator did not appear on the screen at any time. After this pilot run, and on the basis of observation and students' comments, it was decided that this format was too monotonous and did not take advantage of the possibilities of having the teacher appear in person on the television screen from time to time. Therefore, in the presentation of the mathematics program over television, and the subsequent presentation of the English program on television the instructor appeared briefly on the screen to review the content of a previous unit of the program, to introduce and develop a new topic, to summarize, and generally to provide continuity, variety, and a human element.

Preparing the Television Version

The step-by-step procedure employed for modifying a programed unit for televised presentation and the mode of presentation over television were as follows:

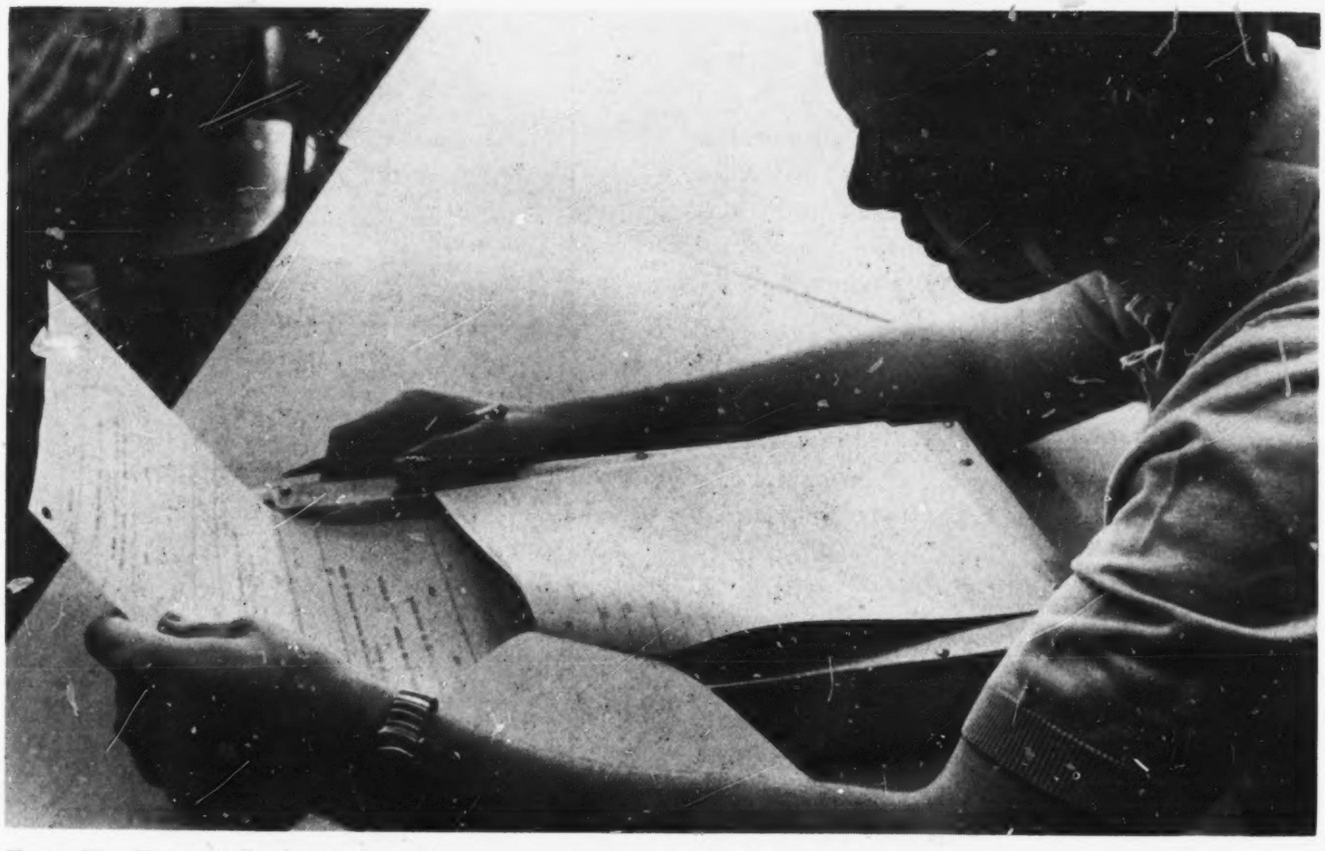


Figure 20. The specially designed teaching machine used in the English experiment.

(a) The programer and television producers carefully read each unit of the program and decided what material should be presented by the instructor on television.

(b) The material to be handled in programed form was then examined, and for each frame a decision was made as to what needed to be shown visually in print form on the television screen and what could advantageously be said through the audio channel.

(c) On the basis of these decisions a script was prepared for each unit which indicated what was to be shown visually for each frame and what would be said by the instructor.

(d) Additional visual materials for the instructor to use on television were prepared by a graphics artist on 8½" x 11" blue paper and each was enclosed in a clear acetate envelope so that the instructor could write on these graphics with a grease pencil during the televised presentation.

(e) The printed material for each frame of the programed sequences was typed on a 5" x 8" card using an IBM electric typewriter with academic bookface type and confining the printed material to an area $2\frac{13}{16}$" wide and $2\frac{3}{16}$" deep. This area permitted the use of a maximum of 26 characters to a line and a maximum of 11 lines per frame. The frames were rarely crowded to the maximum capacity.

(f) The information that was to be spoken by the instructor, either on camera or off-screen, was typed onto the special rolls of paper that are used in a television prompting device, with one lesson to a roll.

Examples of Frame Adaptations

The following examples indicate the adaptation from the original program to the final television format. The original program sequences are given first with examples of adaptation to television following:

A. *Original version of frame:*

15.

Another kind of adverb is formed by the addition of the suffixes *-wise* or *-ward(s)* to nouns.

NOUN		*SUFFIX*		*ADVERB*
length	+	-wise	=	lengthwise
sky	+	-ward	=	skyward
front	+	-ward(s)	=	frontward(s)

In the sentence, "We walked backward," *backward is an* ....

CORRECT ANSWER: 15.
adverb

B. *Televised version of frame:*

VIDEO

Visual #1 15.

NOUN		*SUFFIX*		*ADVERB*
Length	+	-wise	=	lengthwise
sky	+	-ward	=	skyward
front	+	-ward(s)	=	frontward(s)

AUDIO

"Another kind of adverb is formed by the addition of the suffixes *-wise* or *-ward(s) to nouns.*"

Visual #2

In the sentence, "We walked backward," *backward* is an ....

no audio

CORRECT ANSWER:
Visual #3

adverb

no audio

Figure 21.

A. *Original version of frame:*

16.

In the following sentence, supply the adverb formed by the addition of the suffix *-ward* to a noun that would indicate that the class was returning home.

The senior class was .... bound after an enervating trip to the national capital.

CORRECT ANSWER: 16.

homeward

B. *Televised version of frame:*

VIDEO	*AUDIO*
Visual #1 The senior class was .... bound after an enervating trip to the national capital.	"In this sentence, supply the adverb formed by the addition of the suffix *-ward* to a noun that would indicate that the class was returning home."
CORRECT ANSWER: Visual #2 homeward	no audio

Figure 22.

Example of an Introduction to a Unit

The introduction to the unit on adverbs was given by the instructor who appeared on the television screen and made use of certain prepared visual materials.

Below is a portion of the television script indicating the graphics that the instructor used, together with the accompanying audio comments.

INTRODUCTION TO ADVERBS

VIDEO	*AUDIO*
Instructor in person	To date we have treated three classes of words: Class I, the noun or its equivalent; Class II, the verb; and Class III, the adjective. We shall now consider Class IV, the ADVERB.
Visual #1 Adjective + -ly = adverb quick + -ly = —— soft + -ly = —— strange + -ly = ——	I think you know that most adverbs are formed by adding *-ly* to the adjective. (Instructor wrote *quickly*, *softly*, and *strangely* on the graphic with a grease pencil.)
Visual #2 live*ly* adjective or adverb dai*ly* kind*ly*	Most words that end in *-ly* are adverbs. However, some words like *lively*, *daily*, and *kindly* that end in *-ly* can function as adjectives or as adverbs.
Visual #3 AJ 1. She submitted *daily* plans. 2. Do your homework AV *daily*	For example: In sentence #1 *daily* is an adjective. In sentence #2 *daily* is an adverb. We can't say that a word functions as an adverb simply because it ends in *-ly*. Let's go to the program for some practice in forming adverbs and discriminating adjectives from adverbs.

Method of Televising Program

The method of televising the English program closely followed that described for mathematics on pages 34 and 35 of this report.

Two cameras were used. One was focused on the 5" x 8" cards which carried the typed visuals for the programed sequences. (See Figure 23.) The second camera, equipped with a zoom lens, was used to televise the instructor, or, by means of an overhead mirror, to televise the graphic materials used by the instructor from time to time.

The instructor was seated at a table on which was a lectern to hold the graphics. In front of the instructor was the television prompting device containing the narration. Alongside the instructor was an assistant who gave the time cues which indicated a change of frame and operated a chime to alert both the director and student viewers to a change of frame.

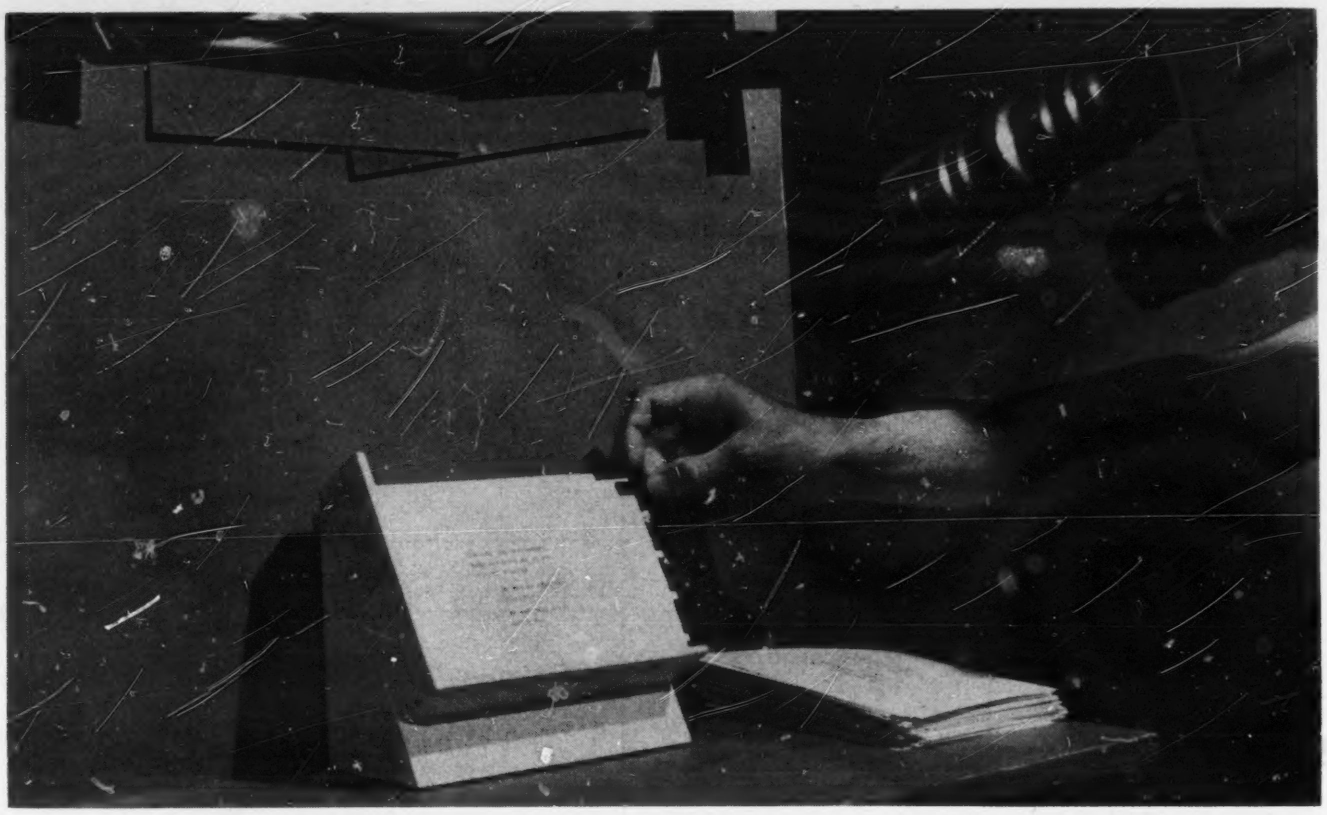


Figure 23. The 5"x8" visual cards arranged before the camera for the television presentation.

On the other side of the instructor was a second assistant who operated the television prompting device. The general arrangement is shown in Figure 24. Both the timekeeper and the director at the controls had copies of the television script showing the visuals and audio comments for each frame. On this script the times were indicated during which each frame would remain on the screen. These times were developed on the basis of a trial run on the teaching machine program during which records were obtained of how long each student took to complete each frame. These records were made using a push button and the multi-channel operations recorder. For the first trial run on television, the mean times for each frame were used, and on the basis of observation of student performance and questioning of the students, these times were accepted as satisfactory. As students become more experienced with this method of learning, these times could probably be reduced. However, in general the students receiving the programed instruction via television indicated that the time allowed for each frame was "about right."

Method of Responding to Programed Sequences Presented via Television

In the television classrooms each student was provided with a special response sheet for each unit.

The response sheet had spaces numbered to correspond with the frames on the television screen. The students were requested to write their responses to each question in the program in the appropriate space, before the correct answer appeared on the television screen. They were informed that this was in their best interest and would increase learning.

For a given unit the sequence of events was generally as follows:

1) The students were given response sheets.

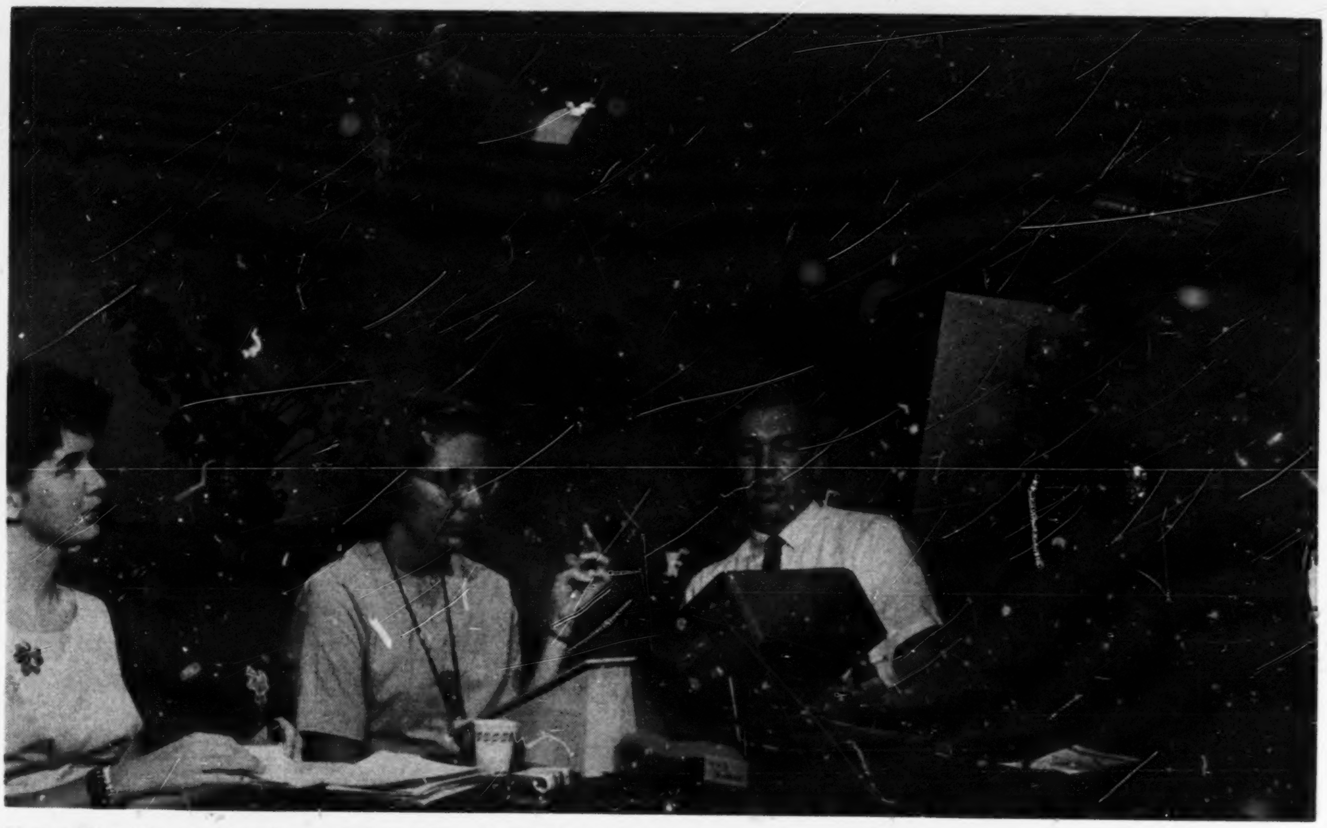


Figure 24. The personnel arrangement for all televised presentations (instructor flanked by two assistants). Note the overhead mirror used to televise the graphic materials.

2) The instructor then appeared on television to review the previous topic and introduce the new unit, using appropriate visual materials.

3) A programed sequence would then be presented over television, and the students would respond to each frame by writing their answers on the response sheet and checking them when the correct answer appeared on the television screen or was heard.

4) The instructor would reappear from time to time on the screen to introduce a new topic, which would be followed by a programed sequence. Finally, the instructor summarized the unit.

5) Immediately after the unit had been completed on television the students took a unit test covering the content of the unit. They were then given a study guide for review of the lesson, and dismissed.

Three studies were conducted which used the programed materials in English. These will be discussed in later sections of this report under the headings of (1) Pilot Project, (2) Experiment I, and (3) Experiment II.

The Tests

by R. E. Spencer

The tests used in the English project fall into three classifications:

1) the regular entrance and placement tests used by The Pennsylvania State University for all entering freshmen

2) English tests specifically designed for the project to assess the degree of learning attained by the subjects immediately following each programed unit and at the end of the course of programed instruction

3) tests measuring areas not directly related to the project under study but used for comparisons of groups; i.e., algebra, arithmetic, School and College Ability Tests, etc.

The tests items themselves fell into four classifications:

1) Items which test specific knowledge of program presentation. These items are actually reworded "frames" from the program, put into a completion or multiple-choice test item format. (Categories 1.11, Knowledge of Terminology; 1.21, Knowledge of Conventions; and 1.23, Knowledge of Classification and Categories in Bloom[1].)
2) Items which may be termed "application" measures of program presentations. Another term might be "usage" items. These items test ability to use correct grammar in a realistic verbal situation. (Category 3.00, *Application* in Bloom.)
3) Items which ask the student to select the "best" of three sentences—each set representing two wrong and one correct method of using a piece of information from the program. (Category 6.00, *Evaluation* in Bloom.)
4) Items which tested areas not directly covered by the program; i.e., vocabulary, paragraph reading, etc., which were not taught by the program but tested an area related to it. (These items may cover several categories in Bloom.)

Table 31 indicates the tests which were used in the experiments involving the program in English grammar.

Reliability coefficients (corrected for attenuation) were computed for the major criterion tests and these are also listed in Table 31. As this experiment used students drawn primarily from the lower range of the distribution (because the project was basic college English), test distributions for the sample were positively skewed on the abilities tests. Tests which were constructed specifically for the sample, however, were more normally distributed, with much greater ranges, and were thus more reliable than the entrance and placement test results.

TABLE 31

TESTS USED IN THE ENGLISH PROGRAM EXPERIMENTS

1. *The Pennsylvania State University Placement Test in English, Form B.* (John S. Bowman, Henry D. Patterson, and Eric V. Sandin.) Composed of 120 items of which 20 are spelling, 50 vocabulary, 25 punctuation, and 25 grammar. Split-half reliability on the Fall 1962 group was .732 (corrected).
2. *Cooperative School and College Ability Tests; College Ability Test, Form 2A.* Cooperative Test Division, Educational Testing Service, Princeton, N. J., 1955. 110 items. Split-half reliability on Fall 1962 group was .859 (corrected).
3. *The Personality Inventory.* Robert G. Bernreuter, 1958. 125 items. Published reliability of dominance subscale was .885.
4. *English Program Final.* Project Staff. 100 items. Split-half reliability on Fall 1962 group was .821 (corrected).
5. *Summed Program Unit-Tests.* Project Staff. 338 items. Kuder-Richardson Formula 21 reliability on Fall 1962 group was .837.
6. *Attitude Test.* Project Staff. 54 items. Saupe's* estimate of the Kuder-Richardson Formula 20 reliability was .644.
7. *The Word Fluency Test.* Project Staff. Subjects were given a three-minute time limit to list as many words as possible that begin with a specific stimulus letter. Saupe's* estimate of the Kuder-Richardson Formula 21 reliability was .658.

*Dressel, Paul L., et al., *Evaluation in Higher Education,* Houghton Mifflin Co., Boston, 1961.

Sub-parts of the major tests used in the validation program were analyzed for use as determiners of specific growth in such areas as grammar, spelling, vocabulary, paragraph reading, punctuation, etc. Reliabilities of these sub-tests were such, however, that they could not be realistically used as valid measures of these sub-areas. The results for all sub-tests were kept through the analyses to make sure that any results would not invalidate the decision to exclude them. It was discovered that they did not add any useful information to the results of the experiments.

---

[1] Bloom, B. S. (ed.), *Taxonomy of Educational Objectives,* David McKay Co., New York, 1956.

The Pilot Project: Fall, 1961

During the Fall Term of 1961, a full-scale pilot project was conducted in the context of the introductory course in English Composition. In this pilot project the entire English grammar program was used in four different formats: programed books, teaching machines, film strips, and closed-circuit television.

In addition to trying out the program, this full-scale trial also provided an opportunity to evaluate and refine the tests and administrative procedures.

As a result of this pilot project, a number of changes were made:

1) As mentioned earlier, it was decided to introduce an instructor from time to time on television to give continuity and to relieve the monotony of presenting only programed material.

2) It was found that, in general, there was insufficient time in a class period (for any of the methods) to include a unit on grammar, a unit on spelling, and a unit-test. (In the subsequent experiments it was decided not to use the spelling units.)

3) The program itself underwent some further refinement based on the measures of learning at the end of the course.

4) The tests were refined and modified to put more emphasis on ability to use correct English grammar and less on ability to define terms.

5) Some tests were dropped from the battery of measures used.

EXPERIMENT V: Comparisons of Televised with Teaching Machine and Televised with Instructor Presentations of English Grammar

by R. E. Spencer

OBJECTIVES

During the summer of 1962, a major experiment was conducted with the programed course in English grammar.

This experiment sought to answer the following questions:

1) Which is more effective in terms of learning—a programed course presented over closed-circuit television to groups of students, or the same material presented to students individually by means of teaching machines?

2) What is the relative effectiveness of the programed course presented over closed-circuit television in comparison with the same content taught in the usual way by an experienced teacher?

3) What are the attitudes of students toward these various methods of instruction?

METHOD

In order to obtain answers to the above questions two experiments were conducted simultaneously—one involving college freshmen, the other involving high school students.

The Populations

The college students who participated in the study were those who scored below 45 on the Penn State English Placement Test. (These represent the lower range of the score distribution.) They were required to take two periods a week of supplemental instruction in English grammar in addition to studying the basic course in English composition.

The high school students were volunteers from four local high schools. They were juniors and seniors, all of whom, accepted the obligation to attend for the full length of the study (10 weeks). They received no other formal instruction during this period.

Treatments

The 43 high school students were randomly assigned to two different treatment groups. One group received programed instruction in English grammar over closed-circuit television; the other group received the same instructional program, but by means of individual teaching machines.

The 30 college students were randomly assigned to two treatment groups. One group received the programed instruction in English grammar by means of closed-circuit television; the other group was taught the same subject matter (not programed), but using the usual lecture-discussion method by an experienced English instructor[1].

The treatments may be summarized as follows:

Comparison A (1) televised programed instruction: 21 high school students
(2) teaching machine instruction: 22 high school students

Comparison B (1) televised programed instruction: 15 college students
(2) face-to-face instruction: 15 college students

All treatments were administered simultaneously, and both groups in the televised treatments received the same instruction but in different television receiving rooms.

Instructional Procedures

The classes met twice a week for a total of 20 meetings. Two days were used at the beginning for pre-testing and two additional days were used at the end for post-tests. Each class period was 75 minutes in length. Approximately 60 minutes of this period were available for instruction. The remaining 15 minutes of each class period were devoted to a unit-test.

All groups received 14 units of instruction; one day was used for orientation, and five days were used for the pre- and post-testing. A careful record of attendance was kept, and no significant differences in attendance were found among any of the treatment groups.

Television Groups

The two television groups received the programed presentations over the television system. Each frame was carefully timed as explained earlier, and all students in these groups advanced at the same rate. The average time per frame was about 30 seconds, but there were variations around the mean depending on the complexity of each frame. The pacing was established in such a way that there was time for every student to respond to every frame. This rate may have caused some students to work at a rate below their capability.

Each student was supplied with a special answer sheet each day, with spaces numbered to conform to the frames in each programed unit of instruction. The students assembled in the television receiving rooms, were handed the special response sheets, and were instructed to respond to each question as it appeared on the television screen and to check their response when the correct response appeared on the screen. Each group was proctored by a graduate student, and in general the students responded actively and continuously during the programed sequences. Figure 25 shows a television class.

Immediately after the completion of a unit of instruction the students took the unit-test, received a study guide for review of the lesson, and were dismissed.

Teaching Machine Treatment

The students receiving the teaching machine treatment were provided with the individual teaching machines which had been especially designed for this research. These machines were loaded with the programed material for

---

[1] James P. Stewart

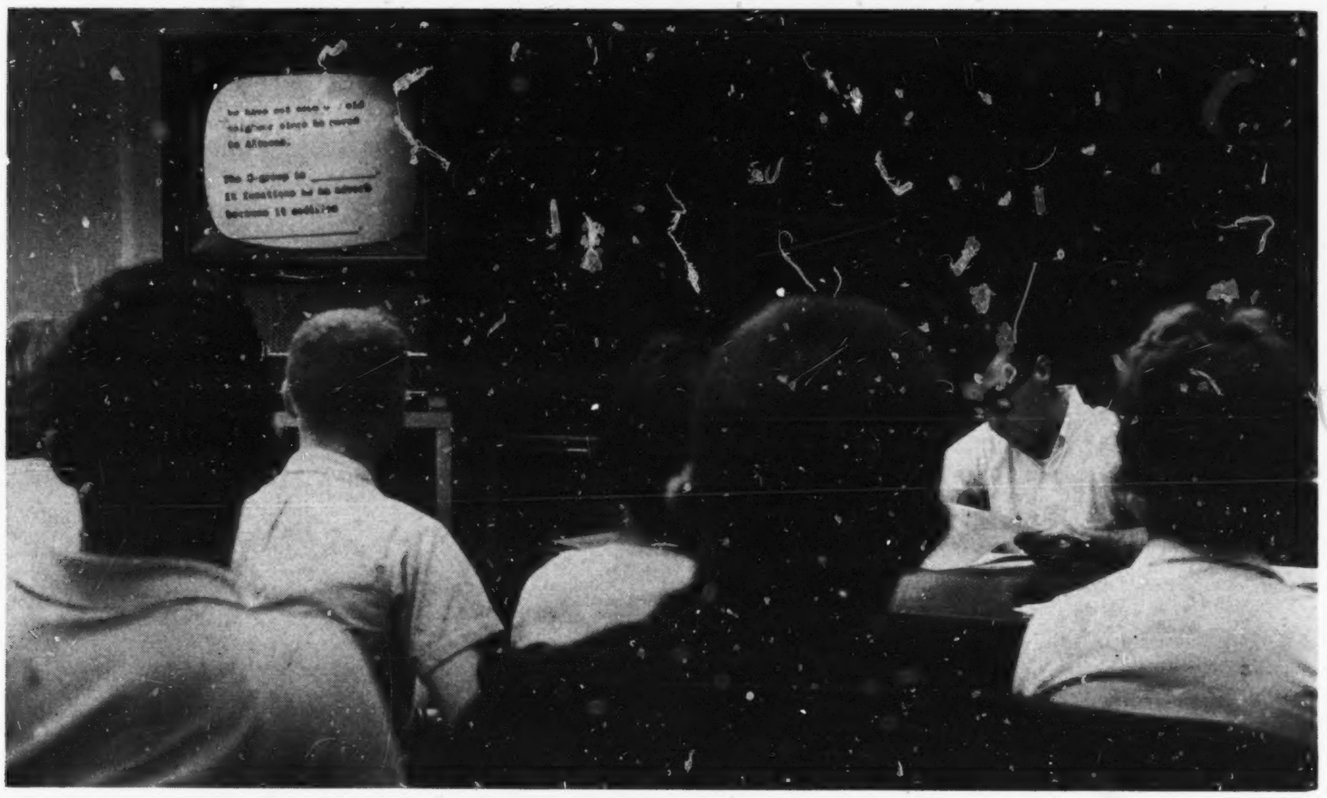


Figure 25. A class session in English as presented over television.

a specified unit of the course, and each student had a copy of the special response sheet. (See Figure 26.)

All students started together at the beginning of each period, and each worked at his own rate until he had completed the unit of instruction. He then was given the unit-test and upon its completion handed it to the proctor, received his study guide, and left the room. The time available for each unit (75 minutes) was adequate for all students to complete the unit.

Face-to-Face Instruction

The group taught by an instructor covered the same topics as the groups receiving the programed treatments. However, the instructor was free to teach as he wished and in general followed a lecture-discussion method.

The unit-tests given each day were the same for both the face-to-face group and the television group. Equal amounts of testing time were also provided.

Testing Procedures

Pre-tests: All students were tested prior to assignment to the various treatments. The college students had already taken the Penn State English Placement Test during the regular college orientation week a few days earlier. The high school students were given the same test.

Post-tests: At the conclusion of the treatments, the English Placement Test was re-administered along with a final examination covering the entire program, and an attitude questionnaire. The sum of each student's scores on the daily unit-tests was also used as a criterion measure.

DIFFERENCES BETWEEN THE ENGLISH AND THE MATHEMATICS EXPERIMENTS

There were some important differences between the college students who participated in the mathematics and English experiments reported here. There were also differences in the way the two courses were handled at the University.

1) Relatively speaking, the population of college students used for the English

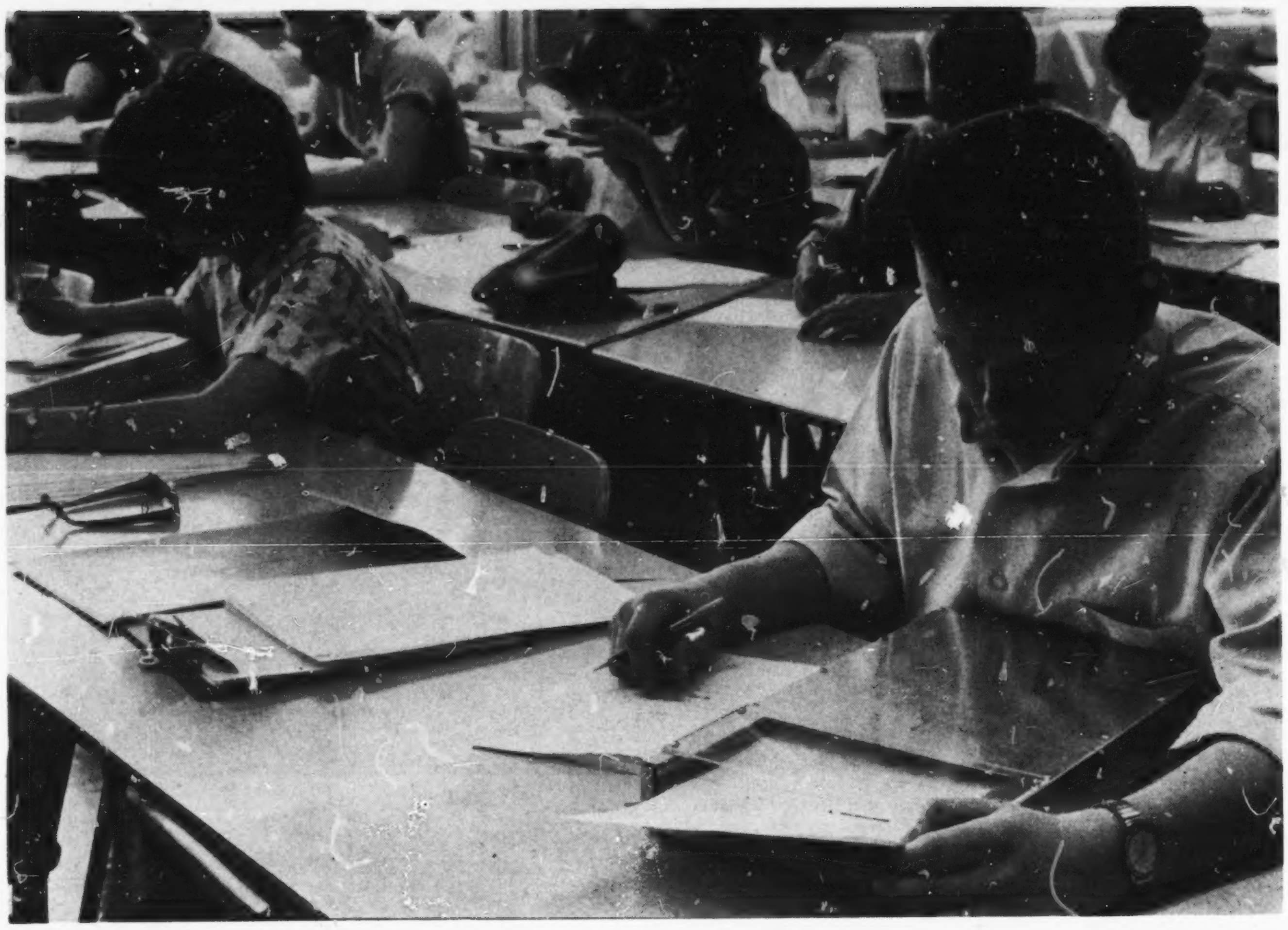


Figure 26. Teaching machine treatment group.

experiments was not as academically advanced in their subject area as those who participated in the mathematics experiments. These students are, in fact, from the lowest part of the distribution of students admitted to Penn State. As the range of abilities is thus restricted, the potentiality of causing variation in learning through two different treatments of instruction is also limited.

2) The programed course in English grammar was conducted as an adjunct to the basic course in English composition for those students who are required to take it. It therefore constitutes an overload of two class periods a week for these students. This may also have limited the learning gains.

3) The English course occupied two class periods a week, whereas the mathematics course required three class periods a week and covered substantially more material.

4) The mathematics course was a regular credit course, whereas the instruction in English grammar and usage did not carry credit, but was a supplement to the basic course in English composition which did, of course, award credit.

The result of these factors may have been to reduce the possibility of obtaining differences between groups in the English experiments. On the other hand, the high school students who participated in the English experiment were volunteers, were well motivated, and were oriented for college. It will be noted that they scored substantially higher on the pre-tests than the college freshmen.

TABLE 32

t-TESTS OF THE DIFFERENCE BETWEEN MEANS FOR THE HIGH SCHOOL TELEVISION GROUP AND THE HIGH SCHOOL TEACHING MACHINE GROUP ON THE PRE-TEST

TEST	Television Group		Teaching Machine Group			
	MEAN	SD	MEAN	SD	t	p
Eng. Placement Total	60.91	17.56	61.23	18.87	.057	$> .90$

Television group N = 21, Teaching machine group N = 22

TABLE 33

t-TESTS OF THE DIFFERENCES BETWEEN PRE- AND POST-TESTS FOR THE HIGH SCHOOL TELEVISION AND HIGH SCHOOL TEACHING MACHINE GROUPS

		PRE		*POST*			
TEST	GROUP	MEAN	SD	MEAN	SD	t	p
Eng. Placement Total	TV	60.91	17.56	65.19	17.82	4.085	$< .01$
	TM	61.23	18.87	63.64	20.38	1.912	.05-.10

TABLE 34

t-TESTS OF THE DIFFERENCE BETWEEN THE HIGH SCHOOL TELEVISION GROUP AND THE HIGH SCHOOL TEACHING MACHINE GROUP ON THE POST-TESTS

	Television Group		Teaching Machine Group			
TESTS	MEAN	SD	MEAN	SD	t	p
Eng. Placement Total	65.19	17.82	63.64	20.38	.260	.70-.80
Program Final Total	73.57	11.47	78.14	11.74	1.259	.20-.30
Daily Unit-Test Total	87.81	6.17	90.36	6.00	1.344	.10-.20
Attitude	25.38	28.83	26.27	24.17	.107	$> .90$

TV N = 21 TM N = 22

RESULTS

Study I: High School Groups

Table 32 presents the t-test comparisons on the pre-test between the high school television and teaching machine groups. No significant difference (at the .05 level) was found on the English Placement Test scores. It can be concluded, then, that the randomization procedure successfully distributed previous achievement in the subject area under experimentation.

Table 33 presents the t-test results on the same test which was given both before and after the instructional program. A significant gain (beyond the .01 level) was observed for the high school television group on the total English Placement Test results. The gains on this test also approached significance for the teaching machine group. Thus, on one of the criterion tests, the Penn State English Placement Test, there is some evidence of a significant learning gain.

Table 34 presents the results for the high school television and teaching machine groups on the various criterion measures given at the end of the treatments.

No significant differences (at the .05 level) were found between the treatment groups on any of these four tests.

Study II: College Group

Table 35 presents results for the comparison of differences in means on the pre-test used for the college television group and the college face-to-face group. No significant difference between the two groups was found, indicating that the randomization procedure successfully distributed previous achievement in the area under experimentation.

Table 36 presents the differences obtained for the measure given both before and after the experimental treatment. Both groups showed significant learning gains on the English Placement Test.

Table 37 presents the results for the groups of college students receiving television or face-to-face treatment on the various criterion measures administered at the end of the course. No significant differences (at the .05 level) were found on the Penn State English Placement Test, the program final, the total of the daily unit-tests, or the grades received by these students in English 1. However, the difference between the groups in attitude toward the methods of instruction was significant beyond the .01 level. This measure indicated that the students receiving face-to-face instruction had more favorable attitudes toward their method of instruction than did the television group.

DISCUSSION OF RESULTS AND CONCLUSIONS FOR HIGH SCHOOL GROUPS

The first question under "objectives" was whether learning under programed instruction would differ if the material were presented over television to a group of students or were presented via teaching machines (to students

TABLE 35

t-TESTS OF THE DIFFERENCE IN MEANS BETWEEN THE COLLEGE TELEVISION AND FACE-TO-FACE GROUPS ON THE PRE-TEST

TEST	Television MEAN	SD	Face-to-Face MEAN	SD	t	p
Eng. Placement Total	43.80	4.37	43.20	5.68	.313	.70-.80

Television group N = 15, Face-to-face group N = 15

TABLE 36

t-TESTS OF THE DIFFERENCES IN MEANS ON PRE- AND POST-TESTS FOR THE COLLEGE TELEVISION AND FACE-TO-FACE GROUPS

TEST	GROUP	*PRE* MEAN	SD	*POST* MEAN	SD	t	p
Eng. Placement Total	TV	43.80	4.37	51.13	6.96	5.013	< .01
	Face-to-Face	43.20	5.68	54.67	8.31	4.926	< .01

Television group N = 15; face-to-face group N = 15

TABLE 37

t-TESTS OF THE DIFFERENCES BETWEEN POST-TEST MEANS FOR THE COLLEGE TELEVISION AND FACE-TO-FACE GROUPS

TESTS	Television MEAN	SD	Face-to-Face MEAN	SD	t	p
Eng. Placement Total	51.13	6.96	54.67	8.31	1.220	.20-.30
Program Final Total	67.53	8.38	70.80	9.87	.945	.30-.40
Daily Unit-Test Total	84.27	4.57	83.73	6.31	.356	.80
Attitude	6.80	27.75	39.07	23.20	3.337	< .01
English I Grades	2.47	0.81	2.53	0.96	.201	.80-.90

Television group N = 15; Face-to-face group N = 15 individually). A group of 43 high school students was randomly divided into a television treatment (N = 21), and a teaching machine treatment (N = 22). Each treatment group received the same programed material and the same criterion measures at the same time. No significant differences in performance were found for the two groups on any of the achievement criterion measures. There were no differences in attitude measures toward the media and methods. Therefore, whether programed instruction is presented over TV or on teaching machines does not result in significant differences in learning as determined by the criterion measures used.

DISCUSSION OF RESULTS AND CONCLUSIONS FOR COLLEGE GROUPS

The second objective of this study was to determine whether there were differences in learning programed material over television in comparison with the conventional face-to-face instructional class covering the same subject matter. Thirty college students were randomly placed into (1) a televised programed (N = 15), or (2) a face-to-face non-programed instructional treatment (N = 15). Each class met for the same length of time, and at the same time of day. A pre-test of the two groups thus randomized showed no significant differences in previous knowledge. Both groups made significant gains in learning as measured by the English Placement Test. However, no differences in performance between the two treatment groups were found on any of the four post-tests which were used to measure learning. Students (of this type) can and do learn as much under programed television presentation as they do under normal face-to-face or conventional instruction, as measured by the several criterion tests used in this experiment.

Significant differences were observed, however, in the expressed student attitudes toward the two treatments. The students in the face-to-face group reported more positive attitudes toward their instructional procedure than did the students in the televised treatment. This difference in attitudes may be attributed (1) to the novelty of the treatment for the TV group, as this was their first experience with programed material over television, and (2) to the personalized interactions of the teacher with the individual students for the face-to-face group.

EXPERIMENT VI: Effects of Personality-Pairing on the Performance of Students in a Programed Course in English Grammar

by Walter Dick & E. L. Seguin

PURPOSE

During the Fall Term, 1962, a further experiment was conducted with students enrolled in the English grammar course. This experiment employed a variation of the paired-student procedure for studying programed materials that was initiated with the Contemporary Algebra program in the Winter Term, 1962 (see Experiment IV, p. 39). In the mathematics study, the students were *randomly* paired and proceeded through the program at a rate determined by both students. With a program placed between them, each member of the pair would construct his own answer to each frame. Incorrect responses were discussed and clarified by the pair before they went on to the next frame. All tests were taken independently; i.e., the members of the pairs did not collaborate in answering the test questions.

In the mathematics study, the students were randomly assigned to pairs as the initial departure from independent study. It was assumed that the results of that experiment would suggest other pairing techniques. Two factors were noted in the mathematics study which appeared to be of some consequence in pairing. These were verbal ability and personality factors. The students indicated on questionnaires that they would like to work with someone with equal "ability." Although the term "ability" was not further defined, the statistical results and observations by the experimenters suggested that verbal ability was an important factor. It was also implicit that personality factors played some part in the compatibility of the pairs. During the course of the mathematics study, for example, the experimenter noted several instances of incompatibility; in fact, there was open hostility in at least one pair. It was this obvious hostility, together with the absence of hostility in other pairs, which prompted this further investigation of the interaction between social structure in the classroom and performance on programed materials, in this instance the English grammar course.

In a discussion of dependence-independence patterns in the classroom situation, Flanders[1] suggests that a student "will learn with greater ease in a given classroom situation, if he can respond in a fashion that is consistent with his previous experience." What happens when this potential for consistency in response is thwarted, when individuals are selectively paired to facilitate or inhibit this consistency? Jensen[2] states that "Students may hold the power to reward or coerce one another, depending upon their ability to frustrate or gratify one another's needs." Since the bulk of the studies on the dynamics of instructional groups indicates that power structure and/or social influence are major contributors to performance and satisfaction with classroom experience, it was decided that selective pairing on the personality dimension of dominance-submission might yield valuable information toward increasing learning.

In addition to investigating the effects of personality traits on performance, the experiment being reported here was designed to evaluate the nature of dyadic interactions when members of the dyad had been paired on personality characteristics; specifically, the selective pairing on dominance mentioned above. It was hypothesized that the incidence of overt conflict situations would be greater in those dyads in which the criterion for pairing was similarity of dominance rather than dissimilarity.

METHOD

Subjects

The subjects were 56 entering freshmen who had failed to achieve the minimum score of 45 on the Penn State English Placement Test. The group was composed of 56 students (50 males and 6 females) from the following curriculums: agriculture, business administration, chemistry and physics, education, and home economics.

Materials

The program on English grammar previously described was used as the basic material to be learned. This was supplemented by a study guide for each lesson. The tests which were utilized are listed in Table 31, p. 56. Briefly, these included the Penn State English Placement Test (EPT), School and College Ability Test (SCAT), Bernreuter Personality Inventory, summed program unit-tests, a word fluency test, an attitude measure, and a program final test. Two additional measures considered in this study were total learning time and English I grades.

An observation record was provided for a classroom observer. This consisted of a checklist of typical behavioral dichotomies, such as whether or not the observed behavior (1) had a task or social orientation, (2) was verbal or motor in nature, and (3) was classified as being a conflict or no-conflict situation. At the end of the checklist there was space to record subjective judgments of the observer and other information which exceeded the scope of the checklist.

Experimental Design

As a first step in the pairing procedure, the students were randomly assigned to two groups of 28 subjects each and then ranked from highest to lowest in terms of dominance scores on the Bernreuter Personality Inventory. (These scores, as well as the initial English Placement Test scores, were obtained from the results of the freshman pre-registration testing.) One group was paired for maximum similarity on the dominance scale and will be identified as the Similar Group. The most dominant student was paired with the second most dominant, etc. (See Figure 27.) The second or Dissimilar Group was paired for minimum similarity between the members of the pair. Therefore, the student ranked number one on dominance in this group was paired with the student ranked fifteenth, then Number 2 with Number 16, etc. (See Figure 27.)

---

[1] Flanders, N. A., "Diagnosing and Utilizing Social Structures in Classroom Learning." In *The Dynamics of Instructional Groups*, National Society for the Study of Education, University of Chicago Press, Chicago, 1960.

[2] Jensen, G., "The Sociopsychological Structure of the Instructional Group." In *The Dynamics of Instructional Groups*, National Society for the Study of Education, University of Chicago Press, Chicago, 1960.

Following an orientation session, all of the students were given the School and College Ability Test. The use of the program on English grammar was initiated during the third session and continued at the rate of two 75-minute classes per week until all 14 units were completed. Two testing sessions followed the completion of the program. The English Placement Test was administered during the first session together with the first half of the attitude questionnaire. During the second session the students were given the word fluency test, the program final, and the second half of the attitude questionnaire. On all of these tests, with the exception of the daily unit-tests, the students worked independently; they received no information from their partners.

PAIRING PROCEDURE FOR THE SIMILAR AND DISSIMILAR GROUPS: STUDENTS RANKED ACCORDING TO SCORE ON BERNREUTER DOMINANCE SCALE

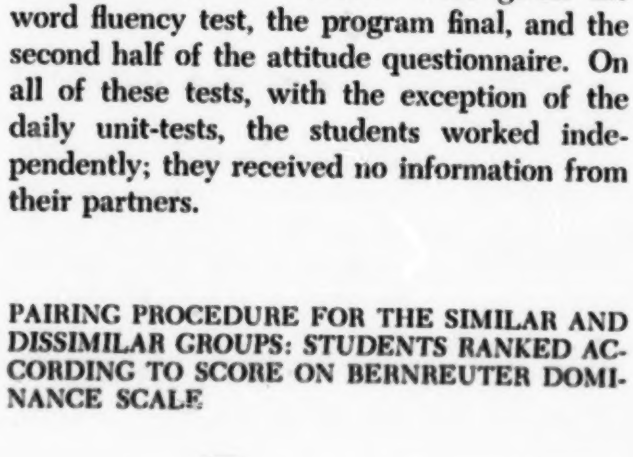


Figure 27.

Classroom Procedure

The learning sessions were conducted according to the following procedure: all students met in the same room and were not aware of belonging to either of the experimental groups. When both members of a pair arrived, they obtained a program unit, two answer sheets, and a 5" x 8" card which was used as a mask. The students seated themselves side-by-side with the program placed between them. (See Figure 28.) They silently read the first frame and *each* constructed his own answer. The mask was then moved down the page and the students' responses were compared with the correct answer. If both students were correct they proceeded to the next frame. However, if one or both of the students made an incorrect response they were instructed to reread the frame and discuss it until both members thought they understood the material.

When the program was completed, the pairs cooperated in answering the items on the daily English usage test. During each of the last six program sessions, the students were independently required to write a short essay on an assigned topic. When this was completed, each student was given a study guide and permitted to leave.

During each class the English unit-test results from the previous lesson were posted for the students. For reasons of control, the tests were never actually returned to the students. However, the six essays were graded and returned.

Make-up sessions were also conducted daily for those students who were unable to attend a regularly scheduled class. At no time was a student permitted to go through a program alone nor could any pair start a unit before they had completed all of the previous units. These controls provided a common sequence of materials and a uniform learning situation for all the students in the experiment.

During each session a qualified psychologist acted as an observer in the classroom. His duty was to observe and note all typical and atypical interactions which took place in the classroom. An interaction has been defined by Thibaut and Kelley[1] as an instance in which

---

[1] Thibaut, J. W., and Kelley, H. H., *The Social Psychology of Groups*, Wiley, New York, 1959, p. 10.

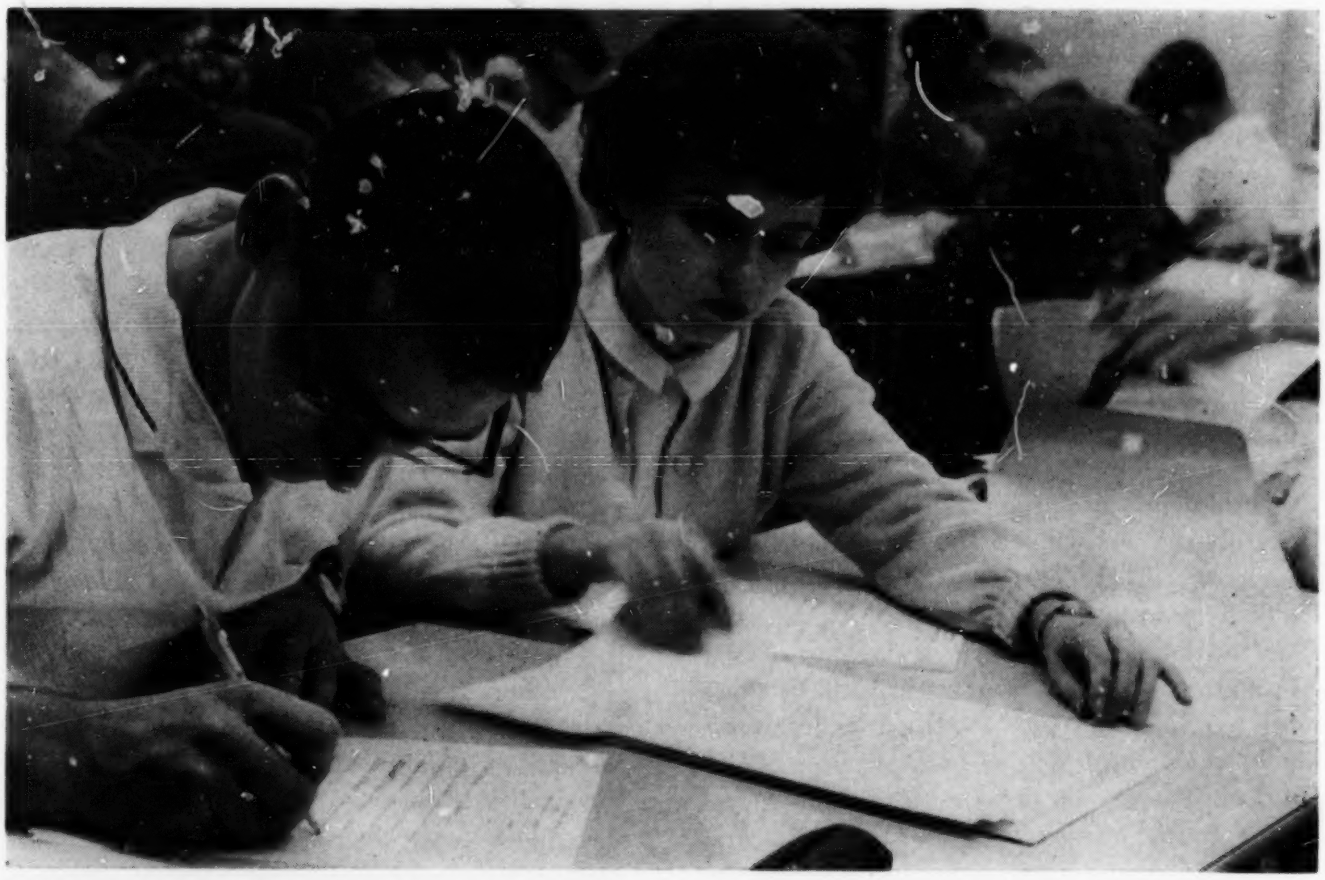


Figure 28. Paired students working on the English program.

there is "... at least the possibility that the actions of each person affect(s) the other." This presented a formidable task for the observer. It was decided, therefore, to adopt the notion of "behavior sequence" or "behavior set" from Thibaut and Kelley.

> Our unit for the analysis of behavior is referred to as the behavior sequence or set. Each unit to be identified consists of a number of specific motor and verbal acts that exhibit some degree of sequential organization directed toward the attainment of some immediate goal or end state[1]

The limitation of having only one observer precluded the idea of observing every pair continuously at every session. Therefore a random schedule of observations was established and maintained throughout the first seven class sessions. At this point, because of examination pre-emptions, absenteeism, make-up sessions, etc., the random observation schedule was abandoned and the observer began to record what, in his opinion, constituted typical behavior.

In order to minimize bias and in an effort to establish a validating criterion for the dominance scale of the Bernreuter Personality Inventory, it was decided that the observer would have no knowledge of (1) where the students ranked on the dominance-submissiveness scale, or (2) the group, (i.e., Similar or Dissimilar) to which the dyads were assigned.

RESULTS

Comparisons of Group Performances

In general, the results seem to indicate no significant differences in performances on the criterion measures between the Similar and Dissimilar groups. Table 38 presents the t-test comparisons between the groups on the pretests. Since no significant differences were found, it can be concluded that the randomization procedure successfully distributed previous achievement in the subject areas under consideration.

---

[1] Thibaut, J. W., and Kelley, H. H., *The Social Psychology of Groups*, Wiley, New York, 1959, p. 11.

TABLE 38

t-TESTS OF THE DIFFERENCES BETWEEN THE SIMILAR AND DISSIMILAR GROUPS ON PRE TESTS

TESTS	Similar MEANS	Similar SD	Dissimilar MEANS	Dissimilar SD	t	p
Dominance Scale	14.61	1.42	14.61	1.70	0.000	> .90
SCAT Verbal	39.25	6.62	38.18	8.13	0.531	.50-.60
SCAT Quantitative	37.75	7.36	36.43	6.06	0.721	.40-.50
SCAT Total	77.00	11.07	74.61	10.18	0.815	.40-.50
English Placement Total-Pre	37.79	5.95	40.50	3.91	1.981	.05-.10

N = 28 for each group; p = .05; t = 2.007

TABLE 39

t-TESTS OF THE DIFFERENCES BETWEEN THE SIMILAR AND DISSIMILAR GROUPS ON POST-TESTS

TESTS	Similar MEANS	Similar SD	Dissimilar MEANS	Dissimilar SD	t	p
Daily Unit Test Total[1]	300.43	11.55	294.57	10.09	1.984	.05-.10
Total Learning Time*	563.64	108.40	545.14	83.06	0.704	.40-.50
Program Final	26.39	6.95	27.57	6.65	0.637	.50-.60
Word Fluency	29.11	6.51	32.36	6.85	1.787	.05-.10
English I Grade	2.14	1.06	1.86	0.95	1.042	.30-.40
Attitude	12.84	3.34	13.00	2.79	0.183	.80-.90
English Placement Total-Post	54.68	7.24	54.68	10.02	0.000	> .90

N = 28 for each group; (p = .05; t = 2.007)

*Time expressed in minutes.

[1] On the daily unit-tests the students worked in pairs; on all other tests they worked as individuals.

Table 39 gives the means, standard deviations, and "t" values for both groups on each of the evaluation measures. In no instance did the differences in the means of the two groups reach the .05 level of significance (p = .05; t = 2.007). Therefore, the null hypothesis cannot be rejected and any differences between the groups must be attributed to chance.

A more refined analysis of the data presented in Table 39 was undertaken to determine if personality subgroups within the two pairing procedures were masking out differences in the overall group performance. The four subgroups whose scores were inspected consisted of dominants who were paired with other dominants (Dom Sim; i.e., dominants who were paired with students with similar dominance ranks); dominants who were paired with submissives (Dom Dis); submissives paired with submissives (Sub Sim); and submissives paired with dominants (Sub Dis). This is shown graphically in Figure 29.

BREAKDOWN OF SIMILAR AND DISSIMILAR GROUPS ACCORDING TO PERSONALITY RANK

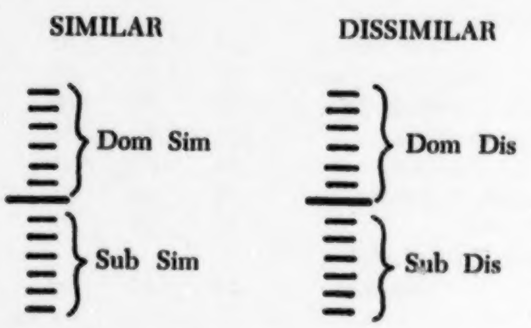


Figure 29.

TABLE 40

t-TESTS OF THE DIFFERENCES IN MEANS FOR SUBGROUPS ON POST TESTS

	DomSim vs. SubSim	DomDis vs. SubDis	DomSim vs. DomDis	SubSim vs. SubDis	DomSim vs. SubDis	SubSim vs. DomDis
Dominance Scale	6.32*	1.62	0.59	0.25	1.61	1.62
Daily Unit Test Total	0.51	0.00	1.35	1.44	1.35	1.44
Total Learning Time	0.91	0.00	0.02	0.94	0.02	0.94
Program Final	0.61	1.80	0.09	0.94	1.73	0.68
Word Fluency	0.31	0.97	0.77	1.62	1.55	0.94
English I Grade	0.35	0.38	0.71	0.72	1.37	1.03
Attitude	2.18*	0.01	1.09	1.20	1.01	1.27
English Placement Total-Post	0.28	1.10	0.66	1.43	0.64	0.90

N = 14 for each group; ($p = .05$; $t = 2.160$)

---

*indicates $p < .05$

TABLE 41

PRODUCT MOMENT r's BETWEEN VARIABLES FOR THE SIMILAR AND DISSIMILAR GROUPS

	Dominance Scale	SCAT Verbal	SCAT Quantitative	SCAT Total	English Placement Total-Pre	Daily Unit-Test Total	Total Learning Time	Program Final	Word Fluency	English I Grade	Attitude	English Placement Total-Post
Dominance Scale		-.01	.15	.10	-.04	.34	.08	.24	.12	.30	*-.53	.17
SCAT Verbal	-.21		.25	*.81	*.48	.17	-.03	.35	-.09	.06	.03	*.62
SCAT Quantitative	.25	.01		*.82	.22	*.45	.18	*.54	.02	.23	.08	*.42
SCAT Total	-.02	*.80	*.60		*.43	*.40	.10	*.57	-.04	.19	.07	*.65
English Placement Total-Pre	.09	*.38	.07	.35		.24	-.08	.31	-.07	.16	.17	*.45
Daily Unit-Test Total	.05	*.40	-.05	.29	*.40		.27	*.43	.04	.26	-.10	.09
Total Learning Time	-.15	.00	-.08	-.04	.03	-.11		*.41	-.13	.05	.03	.12
Program Final	.22	.17	.16	.23	.31	.15	.20		-.14	.33	.02	*.47
Word Fluency	.34	-.01	*.46	.27	*.49	.27	-.20	.32		*.55	.02	.09
English I Grade	.23	.32	.25	*.40	.19	.20	-.14	.10	.32		-.06	*.38
Attitude	-.10	.26	.21	.33	.34	.31	-.08	.30	.15	.18		.06
English Placement Total-Post	.19	*.42	.02	.34	*.63	*.64	-.20	*.47	*.46	.10	*.46	

r's for SIMILAR GROUP

r's for DISSIMILAR GROUP

Note:—r's for similar group located in upper right portion; for dissimilar group in lower left portion.
*$p = .05$; $r = .367$

These four subgroups yielded a total of six possible comparisons of differences in means. The data presented in Table 40 indicate that two of these comparisons reached significance at the .05 level ($p = .05$; $t = 2.160$). In both cases the means were larger for the Dom Sim group. The difference in dominance scale means is an artifact of the experiment in that the only difference between the dominant-similarly paired and submissive-similarly paired groups is the manner in which the members of the two groups responded to the Bernreuter Personality Inventory. The significant difference simply reflects this fact. The other significant difference is in attitude—the Dom Sim group was significantly more favorable toward the course than the Sub Sim.

Relationships Among Variables by Groups

The within-group product-moment correlations for both pre-treatment and criterion measures are presented in Table 41. It should be noted that scores on the verbal portion of the SCAT test did not correlate significantly with either the program final or English I grade for either group. The quantitative SCAT scores did not correlate significantly with these variables for the Dissimilar Group, but they did for the Similar Group on the program final ($r = .54$). The only significant correlation between dominance and the other variables was again in the Similar Group on the attitude measure ($r = -.53$); i.e., within this group, the more "dominant" a student was, the more he tended to have a favorable attitude toward the course.

The significant correlations contained in Table 41 are summarized and presented in Tables 42 and 43. Specifically, these tables indicate those variables which are both: (1) significantly correlated for one group but not for the other, and are also (2) significantly *more* highly correlated (.10 level or better) for the one group than for the other. A less stringent acceptance level ($p = .10$) was deemed appropriate for these tests since one purpose of this study was to generate hypotheses concerning the effects of the differential use of abilities on performance. Consequently the exploratory nature of this segment of the study did not dictate the use of the conventional .05 level.

TABLE 42

VARIABLES BETWEEN WHICH THERE IS A SIGNIFICANTLY HIGHER CORRELATION FOR THE SIMILAR GROUP THAN FOR THE DISSIMILAR GROUP

Variables	*Dz*	*p*
SCAT Total SCAT Quantitative	.464	< .10
SCAT Total Program Final	.414	< .10
SCAT Total English Placement Total-Post	.421	< .10
SCAT Quantitative English Placement Total-Post	.428	< .10
SCAT Quantitative Program Final	.443	< .10
SCAT Quantitative Daily Unit Test Total	.535	< .05
Dominance Scale Attitude	.490	< .05

Note:—Dz represents the difference in z-scores between the two groups.

For example, in Table 42 any pair of variables listed for the Similar Group indicates that the r between these variables was significantly different from zero for the Similar Group, but not significantly different from zero for the Dissimilar Group. Emphasizing this point there is a significantly greater correlation between these two variables for the Similar Group than for the Dissimilar Group.

In order to test for the difference of two r's, a z-transformation was carried out and the difference in z's obtained. Tests were then carried out to determine if this difference in z's was significant. These results suggest that the differential pairing methods resulted in the use of different abilities to achieve comparable performances.

TABLE 43

VARIABLES BETWEEN WHICH THERE IS A SIGNIFICANTLY HIGHER CORRELATION FOR THE DISSIMILAR GROUP THAN FOR THE SIMILAR GROUP

Variables	*Dz*	*p*
English Placement Total-Post Attitude	.437	< .10
English Placement Total-Post Word Fluency	.407	< .10
English Placement Total-Pre Word Fluency	.606	< .02
English Placement Total-Post Daily Unit Test Total	.668	< .01

Note:—Dz represents the difference in z-scores between the two groups.

Analysis of Observational Data

Table 44 shows frequency and percentages of occurrences of behaviors in specified categories for both groups combined and for both groups independently. More specifically, it can be seen that most of the behavior was task-oriented rather than social, more verbal than motor, and that most of the behavior sequences did not result in conflict. The t-tests for proportional differences between the similar and dissimilar groups resulted in no significant differences in any of the specified categories.

DISCUSSION OF RESULTS AND CONCLUSIONS

The question which this study attempted to resolve was whether personality pairing of learners, as defined in this study, contributes differentially to performance on a programed course. More precisely, if pairing does lead to social reinforcement and significantly better learning, can manipulation of the pairing criterion (in this case dominance-submissiveness) indicate a procedure which will yield significantly superior results in performance? Within the framework of this investigation it would seem that there is little or no evidence to support the implied hypothesis.

The purpose of the analysis presented in Table 40 was to show the contribution of the interaction between personality type (either dominant or submissive) and pairing procedure (either similar or dissimilar). Of the 48 possible comparisons, only two were significant. One of these was built into the experiment by the pairing procedure; i.e., in the Similar Group, dominance was *ipso facto* maximized for the Dominant Group and minimized for the Submissive Group. Therefore a significant difference in dominance was created.

The only other significant difference was also obtained between the dominant and submissive similarly-paired groups on attitudes. The Dominant-Similar Group had a significantly more favorable attitude than the Submissive-Similar Group. From an analysis of the means for the other subgroups, it was found that although there were no significant differences between them on attitude, the direction of the differences was consistently in favor of the more dominant group. That is, when a more dominant group was compared with a less dominant group, the more dominant group showed a more favorable attitude toward the programed course in general. It was not surprising, therefore, that a comparison of the extremes on dominance and sub-

TABLE 44

FREQUENCY TABLE OF OBSERVATIONAL DATA SHOWING COMBINED TOTAL AND GROUP PERCENTAGES OF BEHAVIORS WITHIN SPECIFIED CATEGORIES

	Number of Behavior Sequences Observed		Orientation of Behavior				Nature of Behavior				Conflict Observed			
			Task Oriented		Social		Verbal		Motor		Yes		No	
	N	*% of Gp. Tot*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*	*N*	*%*
Similar Pairs	68	100%	67	96%	3	4%	48	62%	30	38%	6	9%	62	91%
Dissimilar Pairs	39	100%	38	95%	2	5%	24	51%	23	49%	8	21%	31	79%
Combined Gp. Totals	107		105		5		72		53		14		93	

mission (when similarly paired) produced a significantly more favorable attitude by the dominants.

This is substantiated by the correlational analysis between dominance and attitude for the Similar and Dissimilar Groups. (See Table 41.) For the former group r = -.53; for the latter, r = -.10. This seems to be consistent with the data presented in Table 40 in that there is a greater relationship between attitude and dominance score when S's are paired similarly.

Although there are no significant differences on the criterion measures, the correlational analysis indicates that perhaps differential abilities are put into use by the students on the various tests. For example, an inspection of the results of the post-administration of the English Placement Test shows that while the two groups had identical means and similar variances (see Table 39), they differed significantly in the correlation on this criterion variable with (1) word fluency, and (2) SCAT Quantitative. It would seem that in the case of the Similar Group, quantitative ability contributes more to performance on the EPT-post than does verbal ability; but for the Dissimilar Group, the reverse is true; i.e., word fluency contributes significantly more to performance than quantitative ability.

The results also indicate no significant difference between the two groups in terms of percentage of the incidence of overt conflict situations. Thus, the hypothesis that the incidence of overt conflict situations would be greater in those dyads in which the criterion for pairing was similarity of dominance, rather than dissimilarity, was not supported.

Several improvements of the observational phase of this study were suggested during the course of the analysis of the data. These were as follows:

1) More than one observer, with adequate training, is needed to increase the reliability of the data.

2) A larger sample of a wider range of student behavior might be valuable to isolate personality characteristics and their social connotations.

3) A larger sample of behavior (more pairs to observe) would also eliminate distortions of the data. For example, one pair contributed seven of the eight conflict situations attributed to dissimilar groups.

It would seem that the results of this study add support to the findings reported by Fishman[1] (in Sanford, 1962) that personality factors do not contribute significantly to the variability in the prediction of academic performance. He states, "The average multiple correlation obtained when aiming the usual predictors (high school grades and scores on a standardized measure of scholastic aptitude) at the usual criterion (freshman average) is approximately .55. The gain in the multiple correlation upon adding a personality test score to one or both of the usual predictors, holding the criterion constant, is usually less than +.05." It may be, however, that personality as a determining factor is "situation specific," and to have any effect it should be relevant to the particular form of ongoing behavior. It is apparent that this study did not involve those modal aspects of behavior which would successfully discriminate the effects on performance of the interaction of personality characteristics and method of pairing subjects on these characteristics.

---

[1] Fishman, Joshua A., "Some Social-psychological Theory for Selecting and Guiding College Students." In *The American College,* (ed. N. Sanford), Wiley, New York, 1962.

PRECEDING PAGE BLANK- NOT FILMED

PART IV

OBSERVATIONS, JUDGMENTS, AND SUGGESTIONS FOR FURTHER RESEARCH

by C. R. CARPENTER

Observations and Judgments

1) Programed content in mathematics and English grammar of adequate quality was organized and prepared in sufficient quantity, with a large enough number of comparable versions, to meet the research requirements.

2) The experiments were conducted with two samples of content which are generally believed to relate to different intellectual abilities—verbal and quantitative.

3) Generally, adequate numbers of student subjects were employed in groups of sufficient size, randomly composed, to meet the requirements of good experimental design.

4) Content was essentially constant for all versions, but formats were necessarily varied for the different media and methods compared.

5) The experiments were conducted in the context of academic curriculums and in learning environments where the usual demands, rewards, and penalties operated, using reliable tests which attempted to sample a wide range of learning performances.

6) The amount and kinds of content and the realistic context of the experiments, the controls and data analysis should permit reasonable generalization of the *principal* results for populations of students represented by the sample used in the research and for the sampled kinds and levels of academic study.

7) A principal result of the research was to show that the programed courses (as developed and administered) can be presented to students for study by the selected media and media combinations with approximately equal probabilities of learning gains.

8) Programed courses of good quality in the content areas sampled can be pre-paced and presented to groups of college and high school students without significantly reducing the learning that occurs under conditions of self-paced and individual study.

9) The main results of the experiments were consistent for the content areas of contemporary algebra and English grammar.

10) Programed content can be adapted and developed for effective presentation over closed-circuit systems when printed forms are used for learner responses and records.

11) Students from the populations sampled have ranges of tolerances for pacing around mean frame-by-frame response cycle rates, and variations within this range do not significantly affect academic learning and attitudes toward programed material.

12) Students of the population sampled can study programed materials in interacting pairs and in classes of different sizes with learning comparable to that which occurs under conditions of individual study. The pairing technique is promising as a means of providing immediate social reinforcement in the study of programed materials.

13) Preliminary experiments suggest that programed courses which are well prepared, carefully administered, and mediated by printed materials (in book or machine form), or by projectuals, film strips, or television, and which are supplemented by study guides and tests will provide conditions as favorable for learning as those found in classes personally taught

by above-average instructors who use the same study guides and tests.

14) The results of the experiments suggest that ranges in achievement scores of students studying programed materials under several conditions are reduced as compared with variations in scores of students taught by good instructors, and that the differences occur mainly in the lower quartile range of the distribution of scores.

15) Students who work in pairs on programed content and who are learning-task oriented show no substantial correlations between a personality measure of dominance-submissiveness and achievement scores.

16) It is possible that the use of supplementary materials such as study guides or assignments, which were common to all groups, may have reduced the possibilities of obtaining differences between treatments. Since the courses being taught by programed materials were parts of the University's regular academic programs, the use of these supplementary exercises was considered necessary for adequate instruction.

Suggested Research and Development Work Based on These Studies

1) Demonstrate the possible extension of the findings on closed-circuit television to instructional broadcast television.

2) Test the generality of results of the media and methods studied by extending research to sound film strips and sound slides, sound motion pictures and video taped recordings of courses, and, for some kinds of instructional content, to audio tapes and radio.

3) Test the generality of findings by extending research to other content areas such as the physical and biological sciences, advanced mathematics, and foreign languages.

4) Test the generality of findings by extending the research to other levels of formal instruction; to the elementary level at one extreme; adult and professional training at the other.

5) Improve the efficiency and economy of programing and producing courses of instruction for different media and integrated combinations of media.

6) Develop improved methods, procedures, and techniques for organizing and producing programed courses of high quality which will exploit the full potentials of sound motion pictures, live television, and video tapes.

7) Conduct research on the functions and forms of print used in combination with non-verbal pictorial and lecture-demonstration materials in programed courses for presentation by means of television, motion pictures, or film strips.

8) Invent and develop practical instrumentation for use with the "new" media for accepting and processing essential response data from students who are working on programed materials, for displaying results to individuals and groups, and for storing and processing the essential information.

9) Improve existing methods for testing the quality and standards of programed materials while courses are being produced.

10) Determine through research how immediate social reinforcement of different kinds can be used with pairs and small groups of students to supplement and summate with the information-response-reinforcement cycle characteristic of programed content when individually administered.